# A Systems-Based Generic Decision Support System Application to Urban Water Management

THÈSE N$^{O}$ 4663 (2010)

PRÉSENTÉE LE 4 MAI 2010

À LA FACULTÉ ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE D'ÉCOHYDROLOGIE
PROGRAMME DOCTORAL EN ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Colin SCHENK

acceptée sur proposition du jury:

Prof. J.-L. Scartezzini, président du jury
Prof. A. Mermoud, Dr M. Soutter, directeurs de thèse
Prof. F. Golay, rapporteur
Prof. R. Mackay, rapporteur
Dr F. Vanderseypen, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2010

# Remerciements

Je remercie en tout premier lieu le docteur Marc Soutter et le professeur André Mermoud, mes deux codirecteurs de thèse. Tous deux, par leur aide, leurs conseils et leur disponibilité, m'ont permis de canaliser mes efforts dans la bonne direction, et ont par ailleurs grandement contribué à la réalisation de ce travail et sa qualité scientifique.

Je remercie également Bastien Roquier, avec qui j'ai eu régulièrement de longues et passionnantes discussions, en particulier sur les aspects systémiques ; bien des éléments de ce travail sont le produit de ces exercices.

Cette étude a eu lieu dans le cadre du programme de recherche européen SWITCH, visant une gestion durable des ressources en eau. Lors des différentes réunions, il m'a été donné d'échanger des idées avec bon nombre de scientifiques actifs dans des domaines similaires et voisins. J'aimerais en particulier remercier mes collègues de notre groupe de travail, Chris Jeffries et Allison Duffy à l'université d'Abertay, Peter van der Steen, Zoran Vojinovic, Solomon Seyoum et Arlex Sanchez à UNESCO-IHE à Delft, Dionysis Assimacopoulos et Elina Manoli à l'université d'Athènes. Mes remerciements particuliers vont enfin au Professeur Rae Mackay pour ses remarques et commentaires constructifs au long se ce travail, et pour son travail de rapporteur de thèse, ainsi qu'à Ewan Last, avec qui nous avons effectué le couplage de son modèle « City Water Balance », et qui a effectué les simulations ; tous deux travaillent à l'université de Birmingham.

D'un point de vue des développements informatiques, la réalisation des diagrammes systémiques n'aurait pu se faire sans les contributions et l'aide de Francois Vanderseypen, de la société Orbifold, que je remercie aussi pour son travail de rapporteur de thèse.

Egalement rapporteur de ce travail de thèse, le Professeur Francois Golay m'a permis, par son analyse et ses remarques constructives, d'améliorer ce travail et de pousser davantage ma réflexion, ce pourquoi je le remercie.

AU niveau des développements et applications, nous avons souvent travaillé en équipe, Marc Soutter, Bastien Roquier et moi-même. Sans leurs contributions, il ne m'aurait pas été possible de pousser aussi loin les concepts et la réalisation de l'outil informatique présenté dans le cadre de ce travail de thèse, et je les remercie donc également pour cela.

Mon travail avait débuté dans le laboratoire HYDRAM, qui est devenu par la suite l'ECHO, et qui a vu se succéder beaucoup de collaborateurs. J'ai apprécié l'ambiance de travail agréable au laboratoire et remercie les nombreuses personnes qui y sont passées ou qui y sont encore, avec qui j'ai eu l'occasion de collaborer, de près ou de loin. En particulier, j'ai passé de très agréables moments avec l'équipe SWITCH : Marc, Bastien et Xavier.

Finalement, ma gratitude va à ma famille en général pour son soutient continu et les bons moments de détente passés, les courses à pied, les voyages : mes parents, mon frère et sa femme, et bien sûr mes enfants et ma femme, Rita.

# Résumé


Dans le domaine de la gestion environnementale, les approches sectorielles, réputées inefficaces, sont encore très répandues, y compris dans la gestion des ressources en eau. Les problèmes de planification et de gestion qui doivent être traités sont dans la plupart des cas complexes, et doivent être appréhendés de manière globale, intégrée et holistique, en prenant en compte les différents domaines de connaissances et les multiples acteurs impliqués. Ainsi, la question du partage et de l'intégration des données, de l'information et des modèles est devenue une problématique cruciale. En ce sens, ces dernières années ont vu le développement de nombreux outils, tels que systèmes d'information, systèmes d'aide à la décision et « frameworks » de modélisation intégrée. Bien que ces outils aient grandement contribué à créer des liens, il n'existe actuellement aucun environnement capable d'intégrer totalement données, modèles, et visualisations. C'est cette problématique que traite la présente thèse, en mettant en œuvre un formalisme sémantique et systémique, avec pour champ d'application principal la gestion de l'eau.

En premier lieu, une description systémique de la gestion de l'eau a été élaborée. Bien que les différents aspects de cette dernière soient abondamment documentés, il manquait une vision intégrée. Ainsi, une représentation graphique des éléments du système et de leurs interconnections, a été réalisée. Le résultat est une référence innovante et transdisciplinaire, une ontologie (une représentation d'un domaine de connaissance) de la gestion de l'eau dans toute sa complexité.

Dans un second temps, un modèle de données a été conçu, sur la base du langage SYSMOD. Ce dernier propose un langage graphique, à la rencontre des ontologies et de la systémique, avec un fort accent sur l'aspect interactionnel. Alors que les modèles de données existants sont en général spécifiques à un domaine ou conçus pour une application donnée, le nouveau modèle de données est générique et transdisciplinaire. Il peut prend en charge tous types de données organisées de manière systémique, et donc potentiellement n'importe quels types de données environnementales.

Pour visualiser et gérer les données systémiques stockées dans la base de données (créée sur la base du modèle de données générique), un nouveau type de visualisation a été développé, proposant des diagrammes de nœuds interconnectés. Alors que les systèmes d'information existants sont souvent des entrepôts de données sans interface utilisateur conviviale, ou des systèmes d'information

géographique centrés sur les données spatiales, cette nouvelle visualisation en diagrammes propose des vues incluant des éléments spatiaux mais aussi non spatiaux (tels que lacs, acteurs ou lois), connectés par des interactions (tels que flux ou influence). Associé à la base de données, l'outil a été intitulé « système d'information systémique » (SIS). Il fournit un moyen de gérer des données systémiques, génériques et transdisciplinaires de manière graphique et intégrée. En outre, son approche systémique facilite l'appréhension et la communication de situations complexes, par leurs représentations en vues thématiques.

En plus de la visualisation en diagrammes, un outil de « reporting » a été développé pour afficher les données liées aux éléments du système (valeurs numériques, commentaires, fichiers, etc.), à la manière d'un navigateur internet. Par ailleurs, basé sur le SIS, un autre module a été créé pour l'interopérabilité avec des modèles externes.

Finalement, une interface utilisateur complète a été créée pour intégrer les développements précédents dans un environnement unique. Pour visualiser les différentes facettes des éléments du système, une vue géographique et un module d'affichage de graphiques ont été ajoutés. L'outil résultant a été baptisé système d'aide à la décision générique (GenDSS), car son application à des situations données permet de créer des vues et liens à des modèles, et donc de fournir des fonctionnalités d'aide à la décision spécifiques à ces différentes situations.

Les développements successifs ont été appliqués à Birmingham (Royaume-Uni), ce qui a permis de montrer que le GenDSS est actuellement limité à la gestion de quelques centaines d'éléments ; au-delà, les temps de chargement et la réactivité de l'interface deviennent inconfortables. D'un autre côté, l'application a aussi démontré les forces de l'outil, qui fournit une solution innovatrice et intégrée pour centraliser les données transdisciplinaires sous différents formats (valeurs numériques, textes, géométries, fichiers), pour décrire de manière holistique des situations grâce aux vues complémentaires, et pour tester différents scénarios à l'aide de modèles liés. Ainsi, le GenDSS s'avère être une solution adéquate et versatile pour de grandes institutions ou entreprises gérant des situations complexes, des données hétérogènes et transdisciplinaires, où les besoins évoluent continuellement, tels que : une industrie automobile, (impliquant plusieurs sites de production et de recherche dans différents pays dotés de lois différentes), un hôpital, ou un réseau de planification stratégique environnementale (pour surmonter les barrières sectorielles et disciplinaires). Dans le cas de la gestion de l'eau, le GenDSS pourrait permettre d'arriver à une réelle gestion intégrée.

**Mots-clés**: données, base de données, modèle, système, information, intégration, eau

# Abstract

In the field of environmental management, inefficient piecemeal sectorial approaches are still widespread, particularly in water resources management, where concerns and decisions are usually split among a set of rather weakly linked stakeholders. Planning and management problems to be addressed are in most cases complex and need to be tackled in a global, integrated and holistic way, including information from the different areas of science and knowledge from the various stakeholders. Therefore, the sharing and integration of data, knowledge and models has become a major issue. A lot of effort has been made in recent years to develop such integrative tools, including information systems (IS), decision support systems (DSS) or integrated modelling frameworks (IMF). Although these tools achieved great progress in terms of linkages, there is currently no environment providing a full integration of the various kinds of data, models, and their visualisation. This is the general issue tackled by the present thesis, with developments based on a semantic, systemic formalism, and applications focusing on water management issues.

The first development step taken was the creation of a systemic description of water management. Albeit the various aspects of water management are abundantly documented and well-known, an integrated description including interconnections was lacking. Therefore, a graphic representation of interconnected system elements annotated with definitions was created. This innovative result constitutes an ontology (a formal representation of a knowledge domain) providing a robust reference and a fairly exhaustive reminder of the complexity of the water management field, including transdisciplinary aspects.

In the second step, a data model was created, based on the new SYSMOD language. The latter proposes a new graphical language uniting the ontological and systemic approaches, with a strong emphasis on interactions. Whereas existing data models are generally specific to a field or designed for a target application, this new data model is generic and transdisciplinary. It can handle any kind of systemically organised data, which potentially includes any sort of environmental data.

To visualise and manage the systems-based, generic data stored within the database (created on the basis of the generic data model) a new kind of visualisation was developed, proposing diagrams of interconnected nodes. Whereas existing information systems are often data warehouses without a

user-friendly graphical user interface, or geographic information systems focussing on spatial data, this new diagram visualisation proposes views including spatial and non-spatial system elements (such as lakes, power plants, stakeholders or laws) interconnected by interactions (such as fluxes or influences). Bound to the database, the tool was called an “information system on the system” (ISS). It provides a way of graphically managing and integrating the systemically organised, generic, transdisciplinary data. Furthermore, the systems approach fosters the apprehension and communication of complex situations in a holistic way (involving many different elements, influences and interactions from various fields), through their representation into thematic views.

In complement to the diagram visualisation, a reporting tool was developed to display the data linked to system elements (numeric values, comments, files, etc.) in a web-browser fashion. To interoperate with external models, another module was developed on the top of the ISS.

Finally, a fully-fledged general user interface was created to integrate all the previous developments into a single environment. To visualise the different facets of system elements, a geographic and a chart view were added, in complement to the reporting view. This final tool was called the generic decision support system (GenDSS), as its application to given situations can give rise to a series of views and linkages to models, thereby providing specific DSS functionalities for these situations.

The successive developments were tested in Birmingham (United Kingdom). Test results indicated that the GenDSS is currently limited to managing a few hundred elements; beyond this quantity, the loading times and the user interface reactivity can become long and unresponsive. On the other hand, the application also demonstrated the strengths of the GenDSS which provides an innovative integrated solution to centralise transdisciplinary data in various formats (e.g. numeric values, texts, geometries, files), to holistically describe situations using complementary visualisations, and to test scenarios using externally-linked models. Thus, the GenDSS may prove an adequately versatile solution for large institutions or enterprises dealing with complex situations, heterogeneous and transdisciplinary data, and where requirements are continuously evolving, such as: a car manufacturing corporation (which would involve several different production and research sites in different countries with different regulations), a hospital, or a strategic environmental planning network (where the tool could help to break the transdisciplinary, sectorial barriers). In the case of water, the GenDSS system could prove very useful to achieve a truly integrated water management approach.

**Keywords:** data, database, model, system, information, integration, water

# Table of Contents

# INTRODUCTION

## DATA AND MODELS INTEGRATION, SYSTEMS, AND DECISION SUPPORT

# Introduction

In the field of environmental management, piecemeal sectorial approaches are still widespread, in spite of the fact that they have been long and widely recognized as not appropriate regarding sustainability. This is notably the case in water resources management, where concerns and decisions are usually split among a set of rather weakly linked stakeholders, taking care in an almost autarkic way of their own specialized area: drinking water supply, waste water disposal, flood control, aquatic ecosystems, energy production, recreational use of water bodies, and so on.

Planning and management problems to be addressed are in most cases complex and therefore need to be tackled in a global, integrated and holistic way. They encompass many different areas of knowledge and science, and involve a bunch of different stakeholders, each of them with its own concerns, background, perspectives, agenda and objectives. It is therefore not surprising that accessibility, sharing and integration of data, knowledge and models have become major issues in contemporary environmental research.

Regarding water management, the arguably most efficient approach is the "integrated water resources management" (IWRM). This methodology calls for various forms of integration at different scales within a general framework. This includes tools integrating data and models, in addition to more organisational instruments, at the legal, institutional or participatory levels. However, IWRM presently remains a somewhat theoretical framework, which still lacks practical, integrative tools.

A lot of effort has been made in recent years to develop such tools, which facilitate this integration, including information systems (IS), decision support systems (DSS) or integrated modelling frameworks (IMF). Although these tools achieved great progresses in terms of linkages, there is currently no environment providing a full integration of the various kinds of data, models, and their visualisation. This is the general issue tackled by the present thesis, with developments based on a semantic, systemic formalism, and applications focussing on the water management issue.

Before describing the general work context, the objectives and the overall layout of this thesis, a brief review of the aforementioned tools and of the general state of the art is provided in order to

set the general background picture in which data integration, semantics and systemic formalism will get their full signification.

## Decision support systems

Decision support systems (DSS) are tools designed to help decision makers analysing a situation or issue, and make well-informed strategic or management choices. Generally, DSSs are software including an information system (IS) and modelling tools. The information system manages and displays the information, and the models provide simulation results. ISs may allow defining different scenarios, simulating future data with the models, and displaying the results to the decision makers.

In the last decades, a very large number of such tools were created. In the year 2009 and for the specific field of water management only, at least eight such various DSSs were proposed and gave rise to a publication (Gastelum and others 2009, Hadihardaja 2009, Halide and others 2009, Kardel and others 2009, Monte and others 2009, Pedras and others 2009, Rao and Rajput 2009, Sulis and others 2009). Generally, DSSs are specifically designed to solve a specific situation, and cannot therefore be easily adapted to other problems and locations; their reuse potential is quite low (Rizzoli and Young 1997).

## Integrated modelling frameworks

To reduce the development times for creating specific DSSs, frameworks exist, which offer suites of reusable components; they are called integrated modelling frameworks (IMFs) (Rizzoli and others 2008). Some of these IMFs propose an application programming interface (API) such as OpenMI (Gregersen and others 2007), TIME (Rahman and others 2003), E2 (Argent and others 2009), or ModCom (Hillyer and others 2001). Using the API, the DSS developer can write code to “wrap” different models within an integrative code which makes the models able to exchange data and run together. Other IMFs propose a dedicated programming language, modelling libraries and a development environment, such as MATLAB (The MathWorks Inc. 2009a) or Mathematica (Wolfram Research Inc. 2009). The modelling libraries contain useful sets of components such as matrices, various calculus algorithms, and visualisation controls (e.g. charts). Using these components, DSSs can be developed. In yet other IMFs, the DSS can be created graphically, within a visual modelling and simulation environment such as STELLA (isee systems Inc. 2009b), Vensim (Ventana Systems Inc. 2009), ithink (isee systems Inc. 2009a) or Simulink (The MathWorks Inc. 2009b). In this case, the developer adds pre-defined visual elements into a diagram and connects them (such as a reservoir

flowing into a water treatment works) into a global model; then, the environment can interpret the global model to run simulations. Finally, some latest developments in IMFs pointed out by Villa and others (2009) adopt a new, ontology-based approach, such as the SEAMLESS project (Janssen and others 2007, van Ittersum and others 2008) -which proposes an integrated assessment framework for agriculture, and the Kepler project (Ludascher and others 2004) -which offers a framework to graphically create scientific workflows (workflows involving heterogeneous data and models). In these cases, an ontological language (see discussion below) is used to provide semantics to the models, their inputs and outputs. Thanks to the semantics, the environment is able to solve interoperability issues such as: what a data is, what kind of data is required by a model, and therefore what data can be provided to a model.

### Information systems

Whereas IMFs' focus is on models and models' data integration, information systems (IS) deal with the integration of information in a larger sense -possibly including comments, documents, pictures, etc.- and the communication of the information. IMFs may include data visualisation facilities (e.g. typically charts), but they are not designed for operations such as storing (in a central database), sharing and displaying pictures or comments by distributed stakeholders through a centralised storage system.

An information system can be defined as an integrated set of components for collecting, storing, processing, and communicating information. In practice, ISs include a multitude of various applications in different fields. In the field of environmental management, the prominent form of IS is probably the geographic information system (GIS) which proposes a maps-based user interface. For instance, several environmental GISs gave rise to publications very recently (Klass and others 2007, Li and others 2007, Nishihama and others 2007, Pradhan and others 2008, Jorge 2009, Shih and others 2009, Soutter and others 2009). Albeit GIS are powerful to display spatial elements, they are more limited when it comes to non-spatial data. Attributes can be linked to spatial elements, but this implies a geo-centrist data structure. Furthermore, GIS are not well suited to represent interconnections (fluxes, causal relations), which makes it difficult to display for example a city's water cycle with its interlinked elements (e.g. treatment works, consumers, wastewater collection).

Another popular, more raw form of information system is the data warehouse. The latter generally presents a simple user interface reflecting the database's fields, and is therefore limited in terms of visualisation (whereas GIS present a rich user interface, generally backed by a kind of database).

There are also many publications relating such developments, including at least six applications to the environmental field (Cooper and others 2008, Comunian and Renard 2009, Kavvadias and others 2009, Miege and others 2009, Peredo-Parada and others 2009, Sander and others 2009).

To sum up, there is a lack of more generic information systems, which would be able to integrate both spatial and non-spatial data, interactions between elements, and various data from different fields in various formats (such as numeric values, texts and multimedia files).

### Data integration and ontological approach

To integrate data, different approaches exist. In software applications, data integration can be achieved at different architectural levels, spanning from a central database to a "manual" integration by the end-users, with intermediate possibilities within applications: in the data access layer, the middleware, the client code, or in the user interface. (Ziegler and Dittrich 2004, Halevy and others 2005, Bernstein and Haas 2008). Such integration traditionally occurs through a schema (a data model) which enforces a fix data structure with rules and constraints. More recently, the ontological approach was proposed as a more versatile, open-world alternative to fix schemas. An ontology, in its modern sense applied to computer science, is a formal structure of knowledge. Ontological formalisms -such as OWL, the language behind the semantic web (Horrocks 2008)- allow defining generic elements such as classes, instances, and properties (e.g. a university student is a subclass of a person; it has the property of being registered in a specific department). Using ontological formalisms, descriptions -semantics- can be applied to the data. For instance, a "4.5" discharge value could be a specific case of discharge value (an instance of the "discharge value" class). And the discharge value class could be a specific case of a numeric value (a subclass of the parent class "numeric value"). The unit "m3/s" could be a specific case of "unit", associated with the "discharge value" class to specify its unit. Using this approach, any kind of data can be described and semantically integrated; the resulting network of interconnected classes and instances is an ontology.

Many field-specific ontologies exist, either to describe only the knowledge related to a field (classes) or to further store the knowledge itself (classes and instances). However, these developments are not turned into fully-fledged information systems with powerful visualisation capabilities (maps, charts, etc.).

### Graphical systems formalisms

A specific case of ontological formalism is the systemic formalism. Systemic formalisms express the knowledge under the specific form of systems: networks of elements (e.g. a lake, or a person, or a law) connected by relations (e.g. a flow or a causal link). Representing a situation with a systemic formalism can help understanding its complexity. The reason why this is important is well illustrated by the beer game example.

The MIT beer game (Sterman 1989) involves at least three teams of players: a retailer, a wholesaler and a beer factory manager. The retailer orders beers from the wholesaler, who in turn orders it from the beer factory. At the beginning of the game, a change in the beer consummation happens at the retailer's shop; this is due to an advertisement campaign. To balance its stock depletion, the retailer orders more beer, but because of the delay between order and delivery, its stock quickly nears complete depletion. In the growing fear of empty shelves, the retailer orders more and more. At the wholesaler's store, this increasing demand is multiplied by the number of retailers. Similarly, the wholesaler gradually increases its demand to balance its stock depletion. At the factory, due to the exploding demand, new workers are hired and the process chain is extended, but this takes time and the orders cannot be fulfilled at first. When the factory finally fully reaches its new production capacity, the demand drops to zero. This is because the retailers' and the wholesaler's stocks are now in excess due to their overreaction to the depletion they experimented. In fact, the consummation by the customers did increase at first, but then quickly stabilised. It did not follow an exponential growth. This game has been played thousands of time in different locations with different people, and it always led to similar results: excess stocks and over production capacity at the factory (Senge (2006)). One main reason of this result is the lack of understanding and knowledge of the whole system behaviour by the different teams. While acting separately, trying to make the best for their own situation, they had no clue on what was occurring elsewhere in the system and did not see the global repercussions of their actions.

Using graphical systems formalism, complex situations such as the beer game crisis can be visually represented into a diagram, including causal links possibly creating retroactions (feedback loops). On that basis, the complexity of the situation can be depicted as a whole (holistically) and therefore better fathomed; interconnections -and sometimes "vicious circles"- appear clearly, which enables stakeholders to take well-informed strategies taking into account the range of possible repercussions (see for instance Senge (2006), Vester (2007) or Meadows (2008) for recent systems thinking references).

**Towards more integration**

In summary, DSSs are tools integrating information and modelling functionalities; they are generally specific to a given problem or set of problems. IMFs provide frameworks to quickly develop DSSs through the integration of models but are limited in terms of information management. Regarding ISs, there is a lack of genericity, in order to integrate and visualise various data from different fields with various formats. For both model and data integration, ontological formalisms propose a promising semantic approach. As a specific case of ontological formalisms, graphical systems formalisms offer a powerful language to visually fathom and communicate complex situations.

Although the aforementioned developments achieved great progresses in terms of linkages, no environment currently provides a full integration of the various kinds of data, models, and their visualisation. This is the general issue that is the focus of this research.

# Thesis framework

This thesis is part of the SWITCH project (Sustainable Water Improves Tomorrows Cities' Health), funded by the European Union. The SWITCH project strives at shifting the current urban water management paradigm, towards a more sustainable approach. It involves many academic teams in many different countries. In Theme 1 (out of six themes), which focuses on strategic approaches, teams in Athens (GR), Birmingham (UK), and Delft (NL) worked on several innovative models. In the Swiss Federal Institute of Technology in Lausanne, CH (EPFL), a platform was to be developed to facilitate the sharing of information, and to link the models created by the partners.

SWITCH involves thirteen test cities all over the world, among which includes Birmingham (United Kingdom), with a population of approximately 5 million inhabitants. Birmingham was selected in priority to apply, test and validate the developments performed at the EPFL. Although theme 1 cooperates closely with a few other cities, the EPFL team tested its developments mainly in Birmingham; in the future, the products will be also implemented in other cities, including Belo Horizonte (Brazil) and Alexandria (Egypt).

This thesis's methodology is in line with the EPFL team's work. Research began with a theoretical study which was then followed by iterative cycles of development of the software and applications, leading to successive improvements. The applications took place mainly on the case of the water management system of Birmingham.

# Thesis objectives

As its main objective, this thesis strives towards a full integration of the various kinds of data, models, and their visualisation, within a single environment. This integration is important, because it provides a generic, centralised way of managing various kinds of data, which are to be used by various kinds of models, and to be visualised by various kinds of views. This integrated environment includes an innovative generic tool addressing the challenge met by enterprises or institutions which face daily, case-by-base, intricate and costly challenges related to sharing and analysing information. To test it, it is to be applied in depth to the field of water management, and it should be submitted to different stakeholders in order to study its usability, and eventually validate it.

The thesis's main objective involves a number of specific research questions:

1. How to systemically define water management?
2. What graphical systemic formalism should be utilized?
3. How to integrate and store various, heterogeneous data?
4. How to visualise and manage the systemic, integrated data?
5. How to leverage the systemic formalism to interoperate with the models?

# Thesis outline

The aforementioned objective and questions gave rise to seven publications (published or submitted). Four of these publications were slightly adapted and correspond to the four chapters of the present thesis. The three other publications, for which the thesis author is second or third author are given in annex.

Question 1 is addressed in *Chapter 1: A water management system model*. Ontological approaches, and more specifically graphical systems formalisms were identified as an interesting, semantic way of defining a situation. Nevertheless, such formalisms had not been applied to the large and intricate field of water management (traditionally discussed in a sectoral way) including its transdisciplinary interconnections. It was therefore necessary to develop such an ontology -a systemic model- as a way to ensure the validity of the systems approach and to provide a reference document for water management applications in the next steps.

The *Chapter 2: A systems-based generic environmental database*, addresses question 3. Existing databases are not designed to generically handle various data. Although ontological approaches offer interesting possibilities, they also present some disadvantages. A mixed approach was taken, leveraging the benefits from both approaches.

Question 4 is tackled in *Chapter 3: A systems-based information system*. Many environmental information systems (IS) are based on a geographic interface (geographic IS -GIS). This is not appropriate to display the systemic information, consisting of nodes and connections. Therefore, a new kind of IS, based on diagrams, was developed.

Finally, the last chapter, *Chapter 4: A systems-based generic decision support system*, presents the final software product, integrating all the various developments, and completing the main objective of the thesis, the provision of an integrated environment to manage and visualise data, and interact with models.

*Annex I: SYSMOD: A systems modelling language for environmental information* addresses question 2, and provides an important theoretical background. Although several graphical systems formalisms

exist, they have not proved adequate regarding the semantic power required to define data, models, and situations. Consequently, a new formalism was created.

*Annex II: ART: An active systems-based navigation and reporting tool* complements Chapter 3 by providing a companion module to display texts, numeric values, and manage files. This visualisation module allocated innovative, dynamic reporting features.

*Annex III: A systems-based integrated modelling framework* addresses question 5. The systemic formalism proposed in Annex I can be used to define data structures and models. Such a formalism made it possible to develop a module to interoperate with models, exporting data towards the latter and importing back the results.

# References

Argent RM, Perraud JM, Rahman JM, Grayson RB, Podger GM (2009) A new approach to water quality modelling and environmental decision support systems. Environmental Modelling & Software 24 (7): 809-818.

Bernstein PA, Haas LM (2008) Information integration in the enterprise. Communications of the Acm 51 (9): 72-79.

Comunian A, Renard P (2009) Introducing wwhypda: a world-wide collaborative hydrogeological parameters database. Hydrogeology Journal 17 (2): 481-489.

Cooper ER, Siewicki TC, Phillips K (2008) Preliminary risk assessment database and risk ranking of pharmaceuticals in the environment. Science of the Total Environment 398 (1-3): 26-33.

Gastelum JR, Valdes JB, Stewart S (2009) A Decision Support System to Improve Water Resources Management in the Conchos Basin. Water Resources Management 23 (8): 1519-1548.

Gregersen JB, Gijsbers PJA, Westen SJP (2007) OpenMI: Open modelling interface. Journal of Hydroinformatics 9 (3): 175-191.

Hadihardaja IK (2009) Decision support system for optimal reservoir operation modeling within sediment deposition control. Water Science and Technology 59 (3): 479-489.

Halevy AY, Ashish N, Bitton D, Carey M, Draper D, Pollock J, Rosenthal A, Sikka V (2005) Enterprise information integration: successes, challenges and controversies. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, ACM: Baltimore, Maryland.

Halide H, Stigebrandt A, Rehbein M, McKinnona AD (2009) Developing a decision support system for sustainable cage aquaculture. Environmental Modelling & Software 24 (6): 694-702.

Hillyer C, Bolte J, van Evert F, Lamaker A (2001) The ModCom modular simulation system. In: (ed) Proceeding of 2nd International Symposium on Modeling Cropping Systems, Florence, Italy, pp 333-343.

Horrocks I (2008) Ontologies and the Semantic Web. Communications of the Acm 51 (12): 58-67.

isee systems Inc. (2009a) iThink - Systems thinking for business. Accessed online 11/17/2009: http://www.iseesystems.com/Softwares/Business/ithinkSoftware.aspx.

isee systems Inc. (2009b) STELLA - Systems thinking for education and research. Accessed online 11/17/2009: http://www.iseesystems.com/softwares/Education/StellaSoftware.aspx.

Janssen S, Ewert F, Li H, Athanasiadis IN, Wien JJF, Therond O, Knapen MJR, Bezlepkina I, Alkan-Olsson J, Rizzoli AE, Belhouchette H, Svensson M, van Ittersum MK (2007) Defining assessment projects and scenarios for policy support: Use of ontology in Integrated Assessment and Modelling.

In: (ed) Proceeding of 2nd International Conference on Asian Simulation and Modelling, Chiang Mai, THAILAND, pp 1491-1500.

Jorge LAB (2009) Soil erosion fragility assessment using an impact model and geographic information system. Scientia Agricola 66 (5): 658-666.

Kardel I, Mirosla-Swiatek D, Chormanski J, Okruszko T, Wassen M (2009) Water management decision support system for Biebrza national park. Environment Protection Engineering 35 (2): 173-180.

Kavvadias KC, Maroulis VZ, Krokida MK, Rahman MS (2009) Development of a Versatile Database for Transport Properties of Foods. International Journal of Food Properties 12 (3): 522-535.

Klass JI, Blanford S, Thomas MB (2007) Use of a geographic information system to explore spatial variation in pathogen virulence and the implications for biological control of locusts and grasshoppers. Agricultural and Forest Entomology 9 (3): 201-208.

Li HY, Shi Z, Cheng JL (2007) ARGIS: an Agricultural Resource Geographic Information System for site-specific management of reclaimable saline soils. New Zealand Journal of Agricultural Research 50 (5): 813-821.

Ludascher B, Altintas I, Berkley C, Higgins D, Jaeger E, Jones M, Lee EA, Tao J, Zhao Y (2004) Scientific workflow management and the Kepler system. In: (ed) Proceeding of GGF Workshop on Workflow in Grid Systems, Berlin, GERMANY, pp 1039-1065.

Meadows DH (2008) Thinking in Systems. Chelsea Green Publishing, White River Junction, Vermont.

Miege C, Choubert JM, Ribeiro L, Eusebe M, Coquery M (2009) Fate of pharmaceuticals and personal care products in wastewater treatment plants - Conception of a database and first results. Environmental Pollution 157 (5): 1721-1726.

Monte L, Brittain JE, Gallego E, Hakanson L, Hofman D, Jimenez A (2009) MOIRA-PLUS: A decision support system for the management of complex fresh water ecosystems contaminated by radionuclides and heavy metals. Computers & Geosciences 35 (5): 880-896.

Nishihama S, Hato T, Sugiura D, Yoshizuka K, Sakurai Y (2007) Development of web environmental geographic information system for evaluating river water environment. Bunseki Kagaku 56 (4): 249-253.

Pedras CMG, Pereira LS, Goncalves JM (2009) MIRRIG: A decision support system for design and evaluation of microirrigation systems. Agricultural Water Management 96 (4): 691-701.

Peredo-Parada M, Martinez-Capel F, Garofano-Gomez V, Atenas M, Riestra F (2009) Eco-hydrological database of chilean rivers: a tool for management of aquatic ecosystem. Gayana 73 (1): 119-129.

Pradhan B, Lee S, Mansor S, Buchroithner M, Jamaluddin N, Khujaimah Z (2008) Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model. Journal of Applied Remote Sensing 2.

Rahman JM, Seaton SP, Perraud JM, Hotham H, Verrelli DI, Coleman JR, Mssanzi (2003) It's TIME for a new environmental modelling framework. In: (ed) Proceeding of International Congress on Modelling and Simulation, Townsville, AUSTRALIA, pp 1727-1732.

Rao BK, Rajput TBS (2009) Decision support system for efficient water management in canal command areas. Current Science 97 (1): 90-98.

Rizzoli AE, Leavesley G, Ascough JC, Argent RM, Athanasiadis IN, Brilhante V, Claeys FHA, David O, Donatelli M, Gijsbers P, Havlik D, Kassahun A, Krause P, Quinn NWT, Scholten H, Sojda RS, Villa F (2008) Integrated modelling frameworks for environmental assessment and decision support. In: Jakeman AJ, Voinov AA, Rizzoli AE and Chen SH (ed) Environmental Modelling, Software and Decision Support, pp 101-118.

Rizzoli AE, Young WJ (1997) Delivering environmental decision support systems: software tools and techniques. Environmental Modelling & Software 12 (2-3): 237-249.

Sander J, Andersen E, Athanasiadis IN, van Ittersum MK (2009) A database for integrated assessment of European agricultural systems. Environmental Science & Policy 12 (5): 573-587.

Senge PM (2006) The Fifth Discipline: The Art & Practice of The Learning Organization (revised edition). Broadway Business, 445 pp.

Shih YC, Chou CL, Chiau WY (2009) Geographic information system applied to measuring benthic environmental impact with chemical measures on mariculture at Penghu Islet in Taiwan. Science of the Total Environment 407 (6): 1824-1833.

Soutter M, Alexandrescu M, Schenk C, Drobot R (2009) Adapting a geographical information system-based water resource management to the needs of the Romanian water authorities. Environmental Science and Pollution Research 16 (S1): 33-41.

Sterman JD (1989) Modeling managerial behavior - misperceptions of feedback in a dynamic decision-making experiment. Management Science 35 (3): 321-339.

Sulis M, Marrocu M, Paniconi C (2009) Conjunctive Use of a Hydrological Model and a Multicriteria Decision Support System for a Case Study on the Caia Catchment, Portugal. Journal of Hydrologic Engineering 14 (2): 141-152.

The MathWorks Inc. (2009a) MATLAB - The Language Of Technical Computing. Accessed online 11/17/2009: http://www.mathworks.com/products/matlab/.

The MathWorks Inc. (2009b) Simulink - Simulation and Model-Based Design. Accessed online 11/17/2009: http://www.mathworks.com/products/simulink/.

van Ittersum MK, Ewert F, Heckelei T, Wery J, Olsson JA, Andersen E, Bezlepkina I, Brouwer F, Donatelli M, Flichman G, Olsson L, Rizzoli AE, van der Wal T, Wien JE, Wolf J (2008) Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96 (1-3): 150-165.

Ventana Systems Inc. (2009) Vensim. Accessed online 17/11/2009: http://www.vensim.com/.

Vester F (2007) The art of interconnected thinking. MCB Verlag, Munchen, 365 pp.

Villa F, Athanasiadis IN, Rizzoli AE (2009) Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. Environmental Modelling & Software 24 (5): 577-587.

Wolfram Research Inc. (2009) Mathematica. Accessed online 11/17/2009: http://www.wolfram.com/products/mathematica/.

Ziegler P, Dittrich KR (2004) Three decades of data integration all problems solved? In: Jacquart R (ed) Proceeding of 18th World Computer Congress, Toulouse, FRANCE, pp 3-12.

# CHAPTER 1

## A SYSTEM MODEL FOR WATER MANAGEMENT

# Preface

This chapter proposes a systemic model for the water management. This model provides a useful reference (an ontology), which is implemented in the next steps into the software as system classes for the case of water management. Subsequently, these derived class elements will be instantiated into specific elements for the Birmingham water management case study.

The systemic model was carried out prior to the definition of the systemic formalism SYSMOD (see Annex I), and therefore makes use of a slightly different language. From chapter 2 onwards, the SYSMOD formalism is used.

This chapter is adapted from the paper: Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management, published in Environmental Management 43 (3): 458-469.

# Abstract

Although generally accepted as a necessary step to improve water management and planning, integrated water resources management (IWRM) methodology doesn't provide a clear definition of what should be integrated. The various water-related issues that IWRM may encompass are well documented in the literature, but they are generally addressed separately. Therefore, water management lacks a holistic, systems-based description, with a special emphasis on the interrelations between issues. This paper presents such a system model for water management, including a graphical representation and textual descriptions of the various water issues, their components and their interactions. This model is seen as an aide-memoire and a generic reference, providing background knowledge helping eliciting actual system definitions, in possible combination with other participatory systems approaches. The applicability of the model is demonstrated through its application to two test case studies.

**Keywords**: integrated water management, water planning, water issues, system model, systems approach, systems thinking

# Introduction

Water management and planning issues are under intense scrutiny, as shown by the abundant related literature and conferences. Many articles call for a new approach to replace the traditional, sectoral way (Niemczynowicz 2000, Postel 2000, Baron and others 2002, Gleick 2003a). The latter emphasizes for instance that if the 20th century's heavy investments in massive infrastructures (dams, aqueducts, centralized treatments, etc.) brought undeniable benefits to billions of people, it also often came along with unexpected social, economical and environmental costs.

The recommended and now commonly accepted methodology (Carter and others 2005) to address water management and planning is the integrated water resources management (IWRM) approach. In many cases, the implementation of IWRM provided encouraging results, such as in New South Wales (Anderson and Iyaduri 2003), where it allowed identifying opportunities that were not previously apparent, as well as in other Australian states (Mitchell 2006), where reductions in the impact of the development on the water cycle were observed. But in general, integrated resources management programs' effectiveness is still difficult to assess; frameworks for evaluation seem to be generally lacking (Bellamy and others 2001). Jeffrey and Geary (2006) also argue that the gap between theory and practice remains extensive, and that the benefits of IWRM have not been clearly demonstrated yet.

An important difficulty regarding the implementation of IWRM is the identification of what to integrate. The probably most quoted definition of IWRM, provided by the Global Water Partnership, states: “IWRM is a process which promotes the coordinated development and management of water, land and related resources, in order to maximize the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems” (GWP 2000). What does “water, land and related resources” exactly imply? In practice, there are a large number of varying interpretations of what to integrate (Biswas 2004), including various blends of components and concepts such as: surface and groundwater, water supply and demand, various uses, different administrative levels, policies, equity, education, health, etc. (Biswas 2004, Mitchell 2006).

The integration of these components and concepts requires understanding the way they interact as a system. Mitchell (2005) proposes two basic interpretations for this system approach: (i)

comprehensive, implying an exhaustive inventory of the variables and relationships, and (ii) integrative, focusing on the identification of key variables and relationships. He mentions that if the former shows generally the drawback of being very time-consuming, the second one may present the risk of overlooking important elements. In parallel, systems approaches may show features of ‘hard’ or ‘soft’ methods (Stephens and Hess 1999), or blends of both. The former strive towards quantification and objectivity, whereas the latter includes qualitative analysis and tends to capture stakeholders’ knowledge through participatory processes.

Recently, a number of promising studies and applications promoted - although they recognise the difficulties - the inclusion of complex, ‘soft’ aspects (linked with some of the social, environmental or economic fields) within water or other natural resources management (Mendoza and Martins 2006, Bosch and others 2007, Collins and others 2007, Pahl-Wostl 2007, Ekasingh and Letcher 2008). The major advantages put forward by these methods are the social learning brought by the process and the gain of a deep understanding of the complex problems tackled. However, one of the drawbacks is the time requirement due to the participatory process, especially when water issues are complex, as the latter may involve numerous components, influences, conflicts and feedback interactions. With this regard, the scientific literature may provide some background help, as it proposes many references about water issues, such as basic and other household needs, industrial, agricultural and energy uses, natural requirements, transportation and recreational purposes, water-related risks, management, planning and governance issues, as well as relations with poverty, health, education, and equity topics issues (Niemczynowicz 2000, Gleick 2003b, Zehnder and others 2003, Falkenmark and others 2004, UNDP 2006, UNESCO-WWAP 2006). However, and although these water-related issues are well documented, they are generally described separately. Therefore, it is difficult to extract the relationships between elements to apprehend the system holistically.

To sum up, the path in IWRM going towards a comprehensive systems approach including ‘soft’ elements can really lead to very rich and interesting insights and robust outputs, but it requires investigation time, for eliciting the components and interactions to be integrated. To facilitate this elicitation process, and therefore reduce the time it necessitates, this paper proposes a generic conceptual model for the water management system. The model synthesises as exhaustively as possible the generic components and interactions involved in water-related issues within a graphical view. It also comes along with an exhaustive review of the water issues, of their related components and, with special emphasis, of their interrelations. This paper also presents two test case studies which demonstrate the applicability and usefulness of the model.

# Modelling methodology

### Inventorying issues

In order to scrutinise exhaustively the water management system, the selected methodology consists in a top-down, issues-components approach, starting with the inventory of the water-related issues. This inventory is provided by reviewing the literature.

In this paper, issues are to be understood in a general way. They embody different stakeholders' various needs, such as freshwater demand and flood protection, as well as ecosystem requirements. They also include governance and management issues, such as water pricing or stakeholder participation rules. Finally, there are also issues beyond the strict field of water, such as for example health and tenure rights problems, or the state of electric infrastructure.

### Inferring the components

While issues are no concrete objects, they however correspond to real components which are their physical response. These components appear generally in the literature about water issues. For instance, the “safe water” issue directly relates to the water resource and to some kind of water supply system infrastructure. Therefore, combining literature information and basic analysis provides the inventory of components.

In this paper, components are concrete objects, classified in two categories: structural elements (such as surface water, sanitation networks, or dams) and non-structural elements (like framework for capacity building, water rights and knowledge database).

### Sketching the system model

Once components are inventoried, the last step consists in organizing them in a clear and logical way, into the system model. As a very important addition, the relationships - structural or functional - must be shown between components during this process. The way to achieve this is a mental exercise of abstraction. The results of such achievements can always be controversial, as there is no unique solution. Different organizations, classifications and generalisations can indeed be imagined, along with different levels of details.

The solution pursued here is a versatile system model. It is meant to enable the addition of further levels of details, and possible new components and new relationships.

# The water management system model

### Graphical output

The developed water management system model is shown in Figure 1. In this drawing:

- Rectangular blocks are components, which belong to one of the two groups: structural (on the right) or non-structural (on the left) elements,
- Arrows terminated with a diamond indicate a specialisation relation (a relation in which “n” given objects are sub-categories of only one hierarchically upper component),
- Arrows terminated by a triangle indicate a functional relation, whose nature is documented by one or a few keywords, and generally by a reference to text descriptions (for instance A12 refers to point 12 of issue A; text descriptions are given in the next section). Arrow-terminated relations might be bidirectional, in which case keywords for both directions are separated by line symbols (———),
- The large grey background rectangle outlines the water-specific domain. As developed in the next section, water-related elements indeed often have connections to elements beyond the strict field of water, such as energy or poverty issues.

### Content overview

The system model shows a relatively dense components and interrelations network, with no particular reading order. In fact, any point is a possible reading entry point, from which it is possible to navigate, following the interrelations. This is further developed in the section regarding the uses of the model through the example of a hydropower plant. To apprehend the model in its whole, seven main water issues are hereafter used as perspectives:

A. Safe water and sanitation
B. Water for agriculture and other food production activities

C. Water for industry, energy and transport
D. Water for recreational, amenity and spiritual purposes
E. Aquatic ecosystems : benefits and pressures
F. Water-related events and hazards
G. Managing and sharing water

In these sections, the coded annotations (such as B2 or G9) refer to the interrelations of the model (Figure 1).

### Safe water and sanitation

Access to freshwater is required for people, for their households and the public buildings. It is necessary for consumption and hygiene purposes, in relevant quantities and qualities. Access to safe basic sanitation is also necessary for people, as a fundamental hygienic, privacy and convenience need.

This issue has obvious impacts on the human health (A1), which in turn influences poverty (A2) issue (for instance through missed days of work). It is also a factor of inequalities, including between genders (A3), when women and girls are in charge of collecting water from distant sources for instance. Sick children and busy girls cannot attend school and therefore this issue also has indirect repercussions on education (A4).

Improving access to water and sanitation might in some cases be achieved through indirect measures. Where people don't feel confident enough or lack the funds for investing into infrastructures for their household (like in slums, where houses have no legitimacy), acting on tenure rights (A5) or providing micro-credits for instance (A6) may prove efficient.

Regarding the legal framework, there are general issues, such as the way institutions are organised, including centralisation or decentralisation trends, which obviously influence the efficiency in the way infrastructures are managed (A7). People may sometimes be directly involved in projects realisation, through construction and maintenance of facilities (A9). Some policies, such as standards (for instance regarding water quality), also play important roles (e.g. for health protection, A8).

Safe water and sanitation infrastructures are of course strongly linked to the water resource, withdrawing it, pouring it back and modifying its quality (A10). They deliver water to and retrieve (and treat) wastewater from different consumers, including private and public buildings, swimming

pools, etc. and provide watering to gardens, sport fields, etc (A12). This possibly also involves storage infrastructures (A11).

### Water for agriculture and other food production activities

To produce food, agriculture needs water. Agricultural productivity (including also non-alimentary production) may particularly be enhanced by irrigation facilities. However, irrigation accounts for about 70 percent of human water use (IFAD 2006). Aquaculture and capture fishing require water as well, with sufficient quantity and quality. Whereas infrastructure is thus used to supply water, waterlogged areas may conversely beneficiate from drainage equipment (B6).

This issue is interrelated with people's health through the occurrence of hunger. As for other health problems, hunger further fosters poverty. Conversely, irrigation or aquaculture facilities may reduce both hunger and poverty in rural areas, directly, through increased food production, but also through secured access to water, employment, and increased area attractiveness (B1). However, depending on land and water tenure rights, certain people (e.g. women) may not be allowed to access, maintain and beneficiate from these facilities, which may therefore create new disparities. Furthermore, they may foster the occurrence of water-borne diseases (B2).

Producing food through irrigation consumes water. Feeding and taking care of livestock further requires water. This leads to the concept of virtual water: producing one kilo of cereals or beef requires a certain amount of water - much more in the latter case. Where water is scarce, importing food with high virtual water content may be an alternative to local production (B3). This also means that food consumption patterns have an influence on the water consumption. For instance, eating less meat would mean a diminution of its production and therefore less water consumed (B4).

Impacts on water consumption in food production may also be induced very indirectly. Although this is controversial, biotechnologies might lead to changes, for instance through the possible finding of organisms resilient to water scarcity (B5).

### Water for industry, energy and transport

Industries need water for product or services generation. It uses this resource in very different ways, such as: constituent part of the product (like beverages), for cleaning, for cooling, to generate steam, etc (virtual water content of produced goods, C4). For transport activities, the need for water is different: the waterways, depending upon their geographic features, may enable ships' circulation.

Water is thus also required to produce energy, as a cooling agent, or for turbine motion in the case of hydropower plants. The latter might further require additional infrastructures, including dams, in the case of large-scale production (C6). Conversely, energy is required by water infrastructures, for instance to provide freshwater (pumping stations), to treat waste water, or to operate devices like dishwashers. Therefore, saving energy may enable water savings and vice versa. This also involves possible savings both ways at the consumers' level (C1). Consumers may also influence water use - and more generally companies' behaviour relative to the environment - through the exertion of pressures (e.g. purchasing eco-products) for environmentally friendly practices (C3).

Industries provide employment and generate production. For operation, they require both water and energy. Therefore, water has an indirect positive influence against poverty. Furthermore, access to reliable electricity provides other opportunities such as powering irrigation systems or enabling activities requiring light after dusk (C2).

Regarding water consumption, reusing wastewater (after a possible treatment) in industries (as well as for irrigation or domestic purposes) offers a technical possibility to save water (but it might require energy). Such measures may be fostered or discouraged by the enforcement framework regulating them directly (reuse standards) or indirectly (water price, subsides, general standards) (C5).

### Water for recreational, amenity and spiritual purposes

Watering sport fields and parks, feeding fountains, swimming pools and spas, providing bathing, sport fishing, navigation, sightseeing and other water-related activities opportunities; these are all uses of water which may play an important social role, providing well-being (D1). Also, possibly multi-purpose objects may provide entertainment opportunities. For instance, dams may provide people with a lake area, for instance for bathing, fishing or windsurfing. All these aspects of water may also be of economic importance through tourism activities (D2).

Regarding enforcement policies, standards are especially important. They may protect human health trough the identification of safe recreational areas, such as beaches showing proper water quality for bathing (D3).

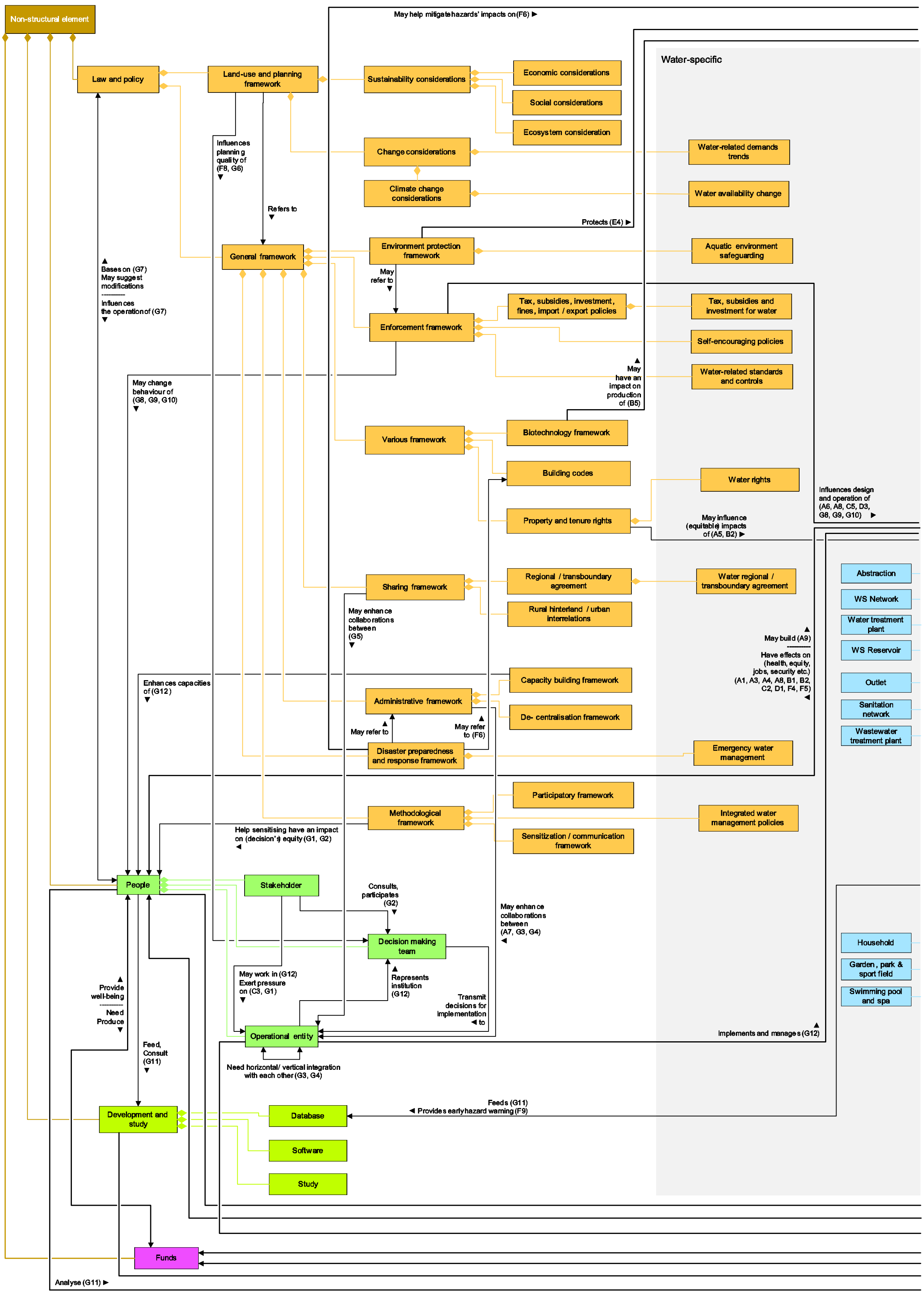



Figure 1: System model for water management (left part)

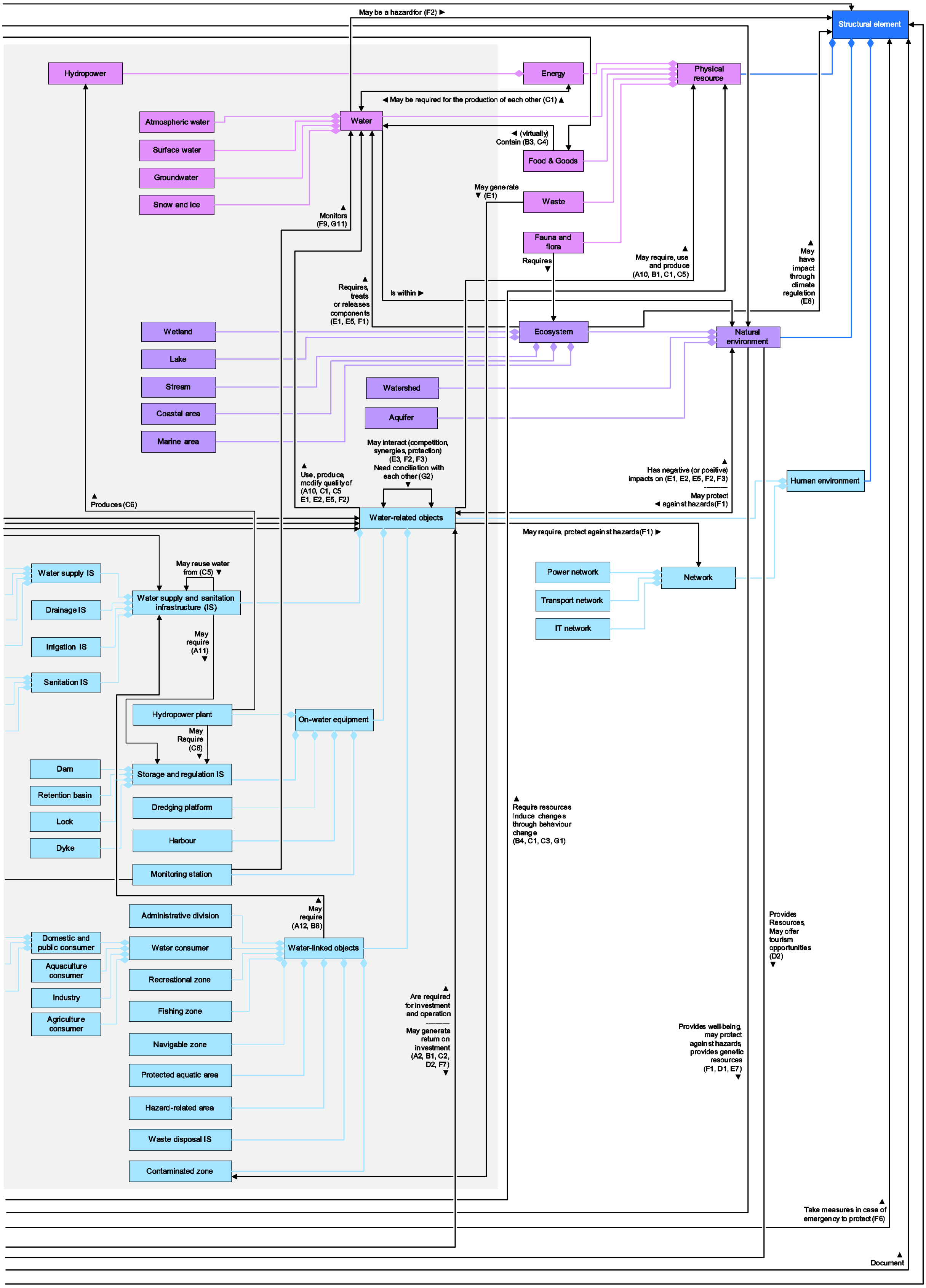



Figure 1: System model for water management (right part)

### Aquatic ecosystems : benefits and pressures

Aquatic ecosystems are required by animal and vegetarian species as habitats. They provide goods and services, needed for socio-economic activities, and they may play a role in risk protection. Conversely, the anthropogenic activities put aquatic ecosystems under pressure, sometimes endangering or even destroying them, along with their animal and vegetarian populations, and thus also along with their goods and services.

Anthropogenic pressures on the aquatic ecosystems are of different types: increased sediment loads, pollutions, flow fragmentation (e.g. dams or locks), invasive species, overuse, etc. Pollutions may occur from point sources (e.g. accidental spills, wastewater outlets) or from diffuse sources (e.g. agricultural fertilisers, soils contaminated by domestic or industrial wastes) (E1). However, certain infrastructures or works may also have positive effects, such as wastewater treatment plants or bank vegetation rehabilitation activities (E2). Moreover, these pressures may also apply to other users, other infrastructures downstream or sharing a common water resource. For instance sediments loads upstream may cause damages or wear to extraction pumps downstream (E3).

Regarding legislation, aquatic ecosystem protection depends upon a wide range of policies and standards, at international, regional and local levels. It also depends upon the way these texts are enforced (E4).

Ecosystems provide a wide range of different benefits. Production aspects are discussed in sections B and C, related to food, industry and energy. Recreational aspects are mentioned in section D, focusing on that topic. The next section, F, covering the hazards related to water, mentions the role played by ecosystems with that regard. Ecosystems also have degradation or dampening capacities of certain components they are exposed to, such as organic loads or sediments. But conversely, such components might be naturally present in some ecosystems and therefore be released into water (E5). At more global levels, ecosystems are involved in the important functions of climate regulation (E6) and genetic resources reservoir (E7).

### Water-related events and hazards

Floods, wave surges, droughts, avalanches are water-related events that may endanger human lives and infrastructures. They may therefore trigger further disasters, like technological hazards. Conversely, other disasters like earthquakes, wars or wind storms may also trigger water emergencies, such as water-lacking refugees' camps, dam breaks, or distribution network failure.

Finally, technological hazards may occur, such as pollutant spills (for instance cases of oil spills caused by foundering vessels), damaging aquatic ecosystems and possibly other water users (F2).

These events may thus heavily damage ecosystems. Conversely, flooding events may contribute to spatial and temporal variability of aquatic ecosystems, and therefore to their richness. Ecosystems may also play an important protection role against hazards (for instance mangrove and coral reef may lessen the strength of waves) (F1).

People may be hit by such events immediately, and also afterwards, due to possible subsequent critical conditions: lack of access to water and water-borne diseases, disruptions of support infrastructures such as roads or health equipments (F4). Poor countries and poor people are generally the most vulnerable to such events, which therefore tend to further increase inequalities (F5).

Regarding protection against water-related hazards, infrastructures may provide mitigation measures: dykes, channels, embankments, retention basins, dams, etc. (F3). Investments in protection may prove beneficial, as it is usually cheaper to invest in protections than to pay for fixing the damages afterwards (Sudmeier-Rieux and others 2006) (F7).

On the other hand, non-structural measures include preparedness, emergency management and response management. Land-use measures, like provision of hazards maps, may prevent constructions in dangerous areas. Building codes may insure good resilience of constructions. Capacity building efforts shall provide managers with the necessary competencies (F6). Monitoring stations, beyond their useful role of general data provision, may also play an important role, when providing early warning information (F9).

Another important aspect regarding the management of water-related events and hazards is the consideration of future changes. In particular, climate change may possibly lead to wetter or drier conditions and therefore to different or increased risks (F8).

### Managing and sharing water

As outlined in the previous sections, water resources are needed for very different uses, including ecosystems safeguarding. These uses may enter into competition, with regard to quantity or quality aspects. Therefore, proper management and sharing are required.

As stated earlier, the presently generally recognised approach for water planning and management is the integrated water resources management (IWRM). This integration shall occur along different complementary axes:

- People: They may be informed and sensitised about water issues, with possible subsequent involvement and behaviour changes. They may be included into the decision process, therefore possibly improving its effectiveness (G1).
- Stakeholders' interests: As water users and related stakeholders may have antagonistic or competitive interests, integrating their views may help leading to solutions taking at best all stakes into account (G2).
- Sectoral institutions (horizontal operational integration): Where different water issues, such as water supply, wastewater treatment, fishing or aquatic systems protection are managed by different institutions, again, integrating their possibly conflict interests may lead to globally best solutions (G3).
- Hierarchical levels (vertical operational integration): When local measures are planned, they shall be in-line with higher levels' strategies, regionally, nationally and internationally (G4).
- Spatial extent: Watersheds and aquifers - if often not coinciding with political boundaries - may represent natural delimitations for water management, as they gather users of a common resource. Therefore, transboundary agreements may play a role for international integration. Moreover, some considerations, like virtual water (import / export) or migrations may involve larger scales (G5).
- Time and sustainability: In order to avoid conflicts between present and future users, time dimension shall be taken into account. It may mean the application of sustainability criteria, integrating the economic, social and ecological aspects (G6).
- Legislative framework: Taking decisions requires an integrated knowledge of the different legal texts and policies, at the international, regional, etc. levels, in order to act concordantly (G7). Then, this legislative framework needs to be enforced. This may be based on strict controls and standards (G8), but also on economic instruments (water pricing, subsidies, fines policies, etc.) (G9) and on self-encouraging incentives (e.g. publications of results, labelling) (G10).
- Data and knowledge: Integrating various kinds of local data is required to provide an information basis on which to rely to take decisions. Furthermore, thanks to access to centralised databases, consulting global knowledge, about successes and failures or best management practices for instance, may help finding solutions (G11). Also, improving local knowledge and local

competencies, through “capacity building”, is important to promote work outputs of good quality (G12).

# Using the model as a general reference

The model, as such, may be used as a reference for better understanding the water management system. For instance, it could be used to answer questions such as: “while implementing a hydropower plant, what components will undergo repercussions?” This example is illustrated in Figure 2, which is a zoom into the system model (Figure1), showing some of the most directly involved components related to hydropower, and their interrelations.

The hydropower plant (1) may require the implementation of storage components, such as dams (2), unless it is a run-of-river plant. It will produce hydropower, which belongs to the energy group (3). As part of the on-water equipment, and of the water-related group, interrelations with other members of these components shall be investigated: for instance, upstream users or activities may release sediments that could on long-term fill the storage unit or wear the turbines, downstream fishing zones may be heavily disrupted, etc. (4). As water is used for the production of energy, its quality may be modified (5) with possible subsequent impacts on the ecosystem (6). The effects of the dam on the latter shall also be investigated: it could provide benefits to the ecosystem through protection against floods, but variability decrease may endanger the habitats (6). Of course, this analysis could be further performed along any other connections and could for instance involve: funds (required for investment), people (they may beneficiate from recreational aspects of a dam lake, but sport fishing or navigation, or beaches activities downstream may suffer from it), operational entities, capacity building and disaster frameworks (availability of specialists able to properly implement and operate the plant, including in emergency cases), methodological frameworks (if not all the stakeholders are represented through a participatory framework, clashes might occur), etc.

This fairly straightforward case - impacts of a hydropower plant - was proposed to illustrate the navigation in the model. However, the latter is designed to help naturally as well within more complex problems, such as: “how could the water supply network’s performance be improved?” or

"how to solve agricultural water needs shortages?" At that level, as shown with the hydropower example, the model provides a way of scanning conceptually - as exhaustively as possible - the space of interrelations and components involved more or less directly in a given issue, possibly pinpointing less intuitive elements that may play a role in it.

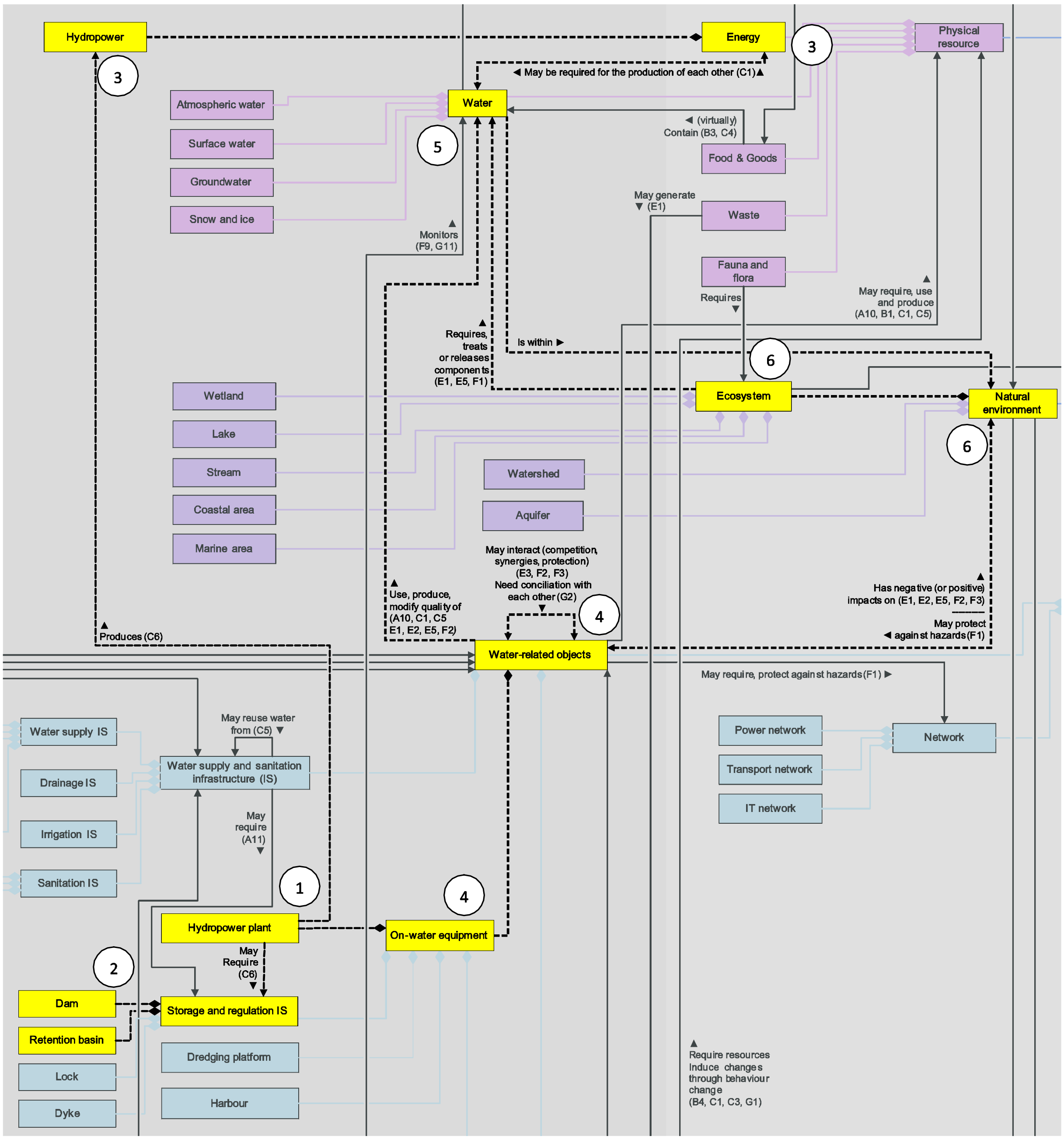



Figure 2: Components and interrelations of the system model which are closely linked to the hydropower plant element

# Test case studies

The model has been applied to two case studies: (i) Birmingham, in England, which accounts with its surrounding municipalities almost 5 million inhabitants, and (ii) Belo Horizonte, in Brazil, accounting around 2.5 million inhabitants. In both cases, the objectives were to test the correctness and the relevancy of the water management system model through:

- The development of a high-level, holistic, overall system representation
- The derivation from the overall view of some usable and concise thematic views

### Application steps

In both cases, the studies were mainly carried out by one person who drove the process. The very first step undertaken by this person was the identification of the stakeholders, through discussions with key people. Then, an iterative process started, consisting in a three-step cycle:

1. Analysis of documents and data
2. Implementation of components and interactions
3. Submission of the developments to the stakeholders for discussion

The analytical and implementation steps made use of the water management system model with the method described in the previous section: the model was used as a reference to exhaustively inventory the system elements. On that generic basis, real components and interactions were derived. To help manage the thereby created voluminous information mass, components and interactions were introduced into a prototype software, which basically allowed storing them - along with some of their features, such as name, icon or colour - in a database and displaying them as diagrams.

The last step was the creation of thematic views, extracting selected information from the overall system network, to provide easier to read, focused views upon a few sub-systems related to chosen water issues.

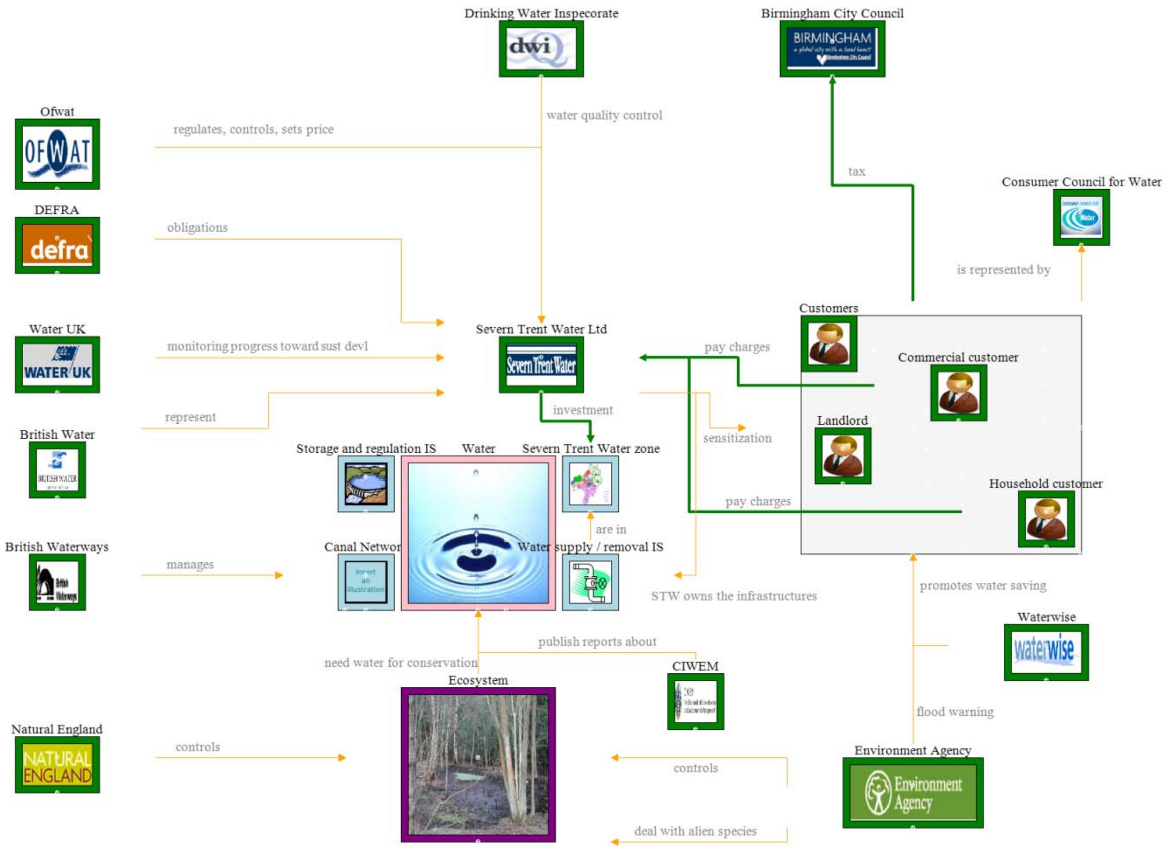



Figure 3: Thematic view: high-level, partial view over the water governance in Birmingham

### Results

In both cases, after an effort of a few months, an overall definition of the water management system has been elicited. In Birmingham, not less than around 700 components and 450 interactions were identified. In Belo Horizonte, around 250 of both were inventoried. The difference can be partially explained by the availability of more information in Birmingham, where a large number of various and detailed studies have been led.

Of course, the diagrams showing all these elements are not directly usable as a communication tool. Thematic views were then created to provide readable outputs, focusing on a given topic and therefore showing only the most relevant, linked information, extracted from the overall system. Such a thematic view, proposed in Birmingham to illustrate governance issues, is shown in Figure 3. This view shows, at a very high level, the stakeholders' responsibilities regarding the water as a resource (central components). This is an example of a very global, partial diagram: it was arbitrarily

chosen to display only a small number of components and relationships in aid of clarity; the overall systemic view contains indeed much more details vertically (e.g. components, sub-components, etc. within the ecosystem or the water-linked infrastructures) and horizontally (e.g. relationships with the laws and policies or with studies). Distinction was made between the influences relationships (light thin connectors) and the money fluxes (darker thick connectors).

The developments of the holistic overall system and of the thematic views provide a very rich system-related knowledge base for further developments. They hereby also serve the purpose of demonstrating the applicability of the water management system model as a generic reference.

# Discussion

The overall, holistic water system definition in Birmingham and Belo Horizonte was produced mainly on the basis of one person's leadership, who analysed the situation, proposed implementations and submitted them to the stakeholders in an iterative manner. However, more participatory application steps, for example through stakeholders' platform meetings, could be followed as well. In that sense, the hereby proposed generic model approach is not mutually exclusive with other systems-based approaches. It is rather considered as a tool facilitating the process by providing background knowledge.

There are many methods and tools used in systems-based approaches used to capture stakeholders' knowledge through participatory approaches. For instance, Bosch and others (2007) present cases where stakeholders were involved in systems thinking, using techniques and tools such as influence diagrams construction, blackboard-supported modelling, participatory matrixes elaboration and computer-based modelling. Pahl-Wostl and Hare (2004) applied different techniques in a participatory approach, including mental models building, hexagon modelling and card sorting to elicit respectively individual representations, system data and actors' network. Collins and others (2007) applied a systems approach through an iterative process which involved the formulation of systems of interest to take into account the multiple stakeholders' perspectives.

The water management system model, since it proposes a generic structure, might therefore be used within such other systems-based approaches applied to water management. In any case, be it a more expert-oriented or a more participatory system elicitation approach, the water management system model may arguably bring important benefits. First, it may reduce the time requirement for the system elicitation process, through the provision of generic background knowledge. This is especially true on longer usage terms: once a system is described holistically, for instance in a city, it may act as a reusable source of information. Indeed, thematic views, focusing on any topic of interest, may be extracted from this repository. Also, as it strives to be as exhaustive as possible, the model provides an aide-memoire helping not to forget possibly important interrelations. This advantage may however present a backside if the model is to be followed slavishly, as some unusual interactions - not inventoried in the model - might be overlooked; therefore, this point should always be considered in the discussion and submission processes with the stakeholders. Finally, and as mentioned earlier, one of the major advantages of participatory processes is the social learning it brings. If the water management system model is to be used in a more expert-oriented approach, this is an aspect that should be taken into consideration.

Although the water management system model would beneficiate from further applications at other scales and in other contexts than large cities, the experiences in Birmingham and Belo Horizonte gave encouraging results, with the creation of rich, reusable and holistic systems definitions, along with some related thematic views.

# Conclusion and perspectives

In this study, the issues related to water are reviewed as exhaustively as possible, their related components are inferred, and their interrelations are emphasised. This analysis allows subsequent sketching of the elements into a general model for water management. The latter was applied to two test case studies, two large cities. It enabled, as a generic reference, to derive many components and interrelations, in order to define their overall water system. From this holistic definition, thematic

views were proposed, extracting selected information to provide more readable displays, focusing on given topics.

Whereas water-related issues are well-known, the analysis into components and interrelations for the realisation of a holistic graphical model is an original contribution. This model can be used as an aide-memoire, a generic reference. It can be applied in conjunction with other systems-based, participatory methods, or in a very expert-oriented manner, allowing reductions of time requirements for systemic analysis. This might be particularly relevant in situations where an overall system definition is required, as in such cases, the very rich holistic repository of inventoried components and interactions might be reused. In conclusion, this model may help taking the path of a comprehensive systems approach, including 'soft' system elements, in integrated water management.

Regarding the perspectives, an important aspect emphasised by the test case studies, is the need for an advanced tool helping managing the complex and numerous system-related information. This software may not only allow storing and displaying components and interactions, but also dealing with related data, such as numeric values, comments, problem notifications, etc. It could therefore lead to the realisation of an information system dedicated to the management of systems-based data, an "information system on the system". Such an information system may enable the creation of interesting advanced thematic views featuring for instance proportional fluxes arrows and problems structures views.

# References


Anderson J, Iyaduri R (2003) Integrated urban water planning: big picture planning is good for the wallet and the environment. Water Science and Technology 47 (7-8): 19--23.

Baron JS, Poff NL, Angermeier PL, Dahm CN, Gleick PH, Hairston NG, Jackson RB, Johnston CA, Richter BD, Steinman AD (2002) Meeting ecological and societal needs for freshwater. Ecological Applications 12 (5): 1247--1260.

Bellamy JA, Walker DH, McDonald GT, Syme GJ (2001) A systems approach to the evaluation of natural resource management initiatives. Journal of Environmental Management 63 (4): 407--423.

Biswas AK (2004) Integrated water resources management: A reassessment - A water forum contribution. Water International 29 (2): 248--256.

Bosch OJH, King CA, Herbohn JL, Russell IW, Smith CS (2007) Getting the big picture in natural resource management - Systems thinking as 'method' for scientists, policy makers and other stakeholders. Systems Research and Behavioral Science 24 (2): 217-232.

Carter N, Kreutzwiser RD, de Loe RC (2005) Closing the circle: linking land use planning and water management at the local level. Land Use Policy 22 (2): 115--127.

Collins K, Blackmore C, Morris D, Watson D (2007) A systemic approach to managing multiple perspectives and stakeholding in water catchments: some findings from three UK case studies. Environmental Science & Policy 10 (6): 564-574.

Ekasingh B, Letcher RA (2008) Successes and failures to embed socioeconomic dimensions in integrated natural resource management modeling: Lessons from Thailand. Mathematics and Computers in Simulation 78 (2-3): 137-145.

Falkenmark M, Gottschalk L, Lundqvist J, Wouters P (2004) Towards integrated catchment management: Increasing the dialogue between scientists, policy-makers and stakeholders. International Journal of Water Resources Development 20 (3): 297--309.

Gleick PH (2003a) Global freshwater resources: Soft-path solutions for the 21st century. Science 302 (5650): 1524--1528.

Gleick PH (2003b) Water use. Annual Review of Environment and Resources 28: 275--314.

GWP (2000) Integrated Water Resources Management. Global Water Partnership -Technical Advisory Committee, Stockholm, Sweden, 67 pp.

IFAD (2006) Water for Food, Agriculture and Rural Livelihoods. In: UNESCO-WWAP (ed) Water, a shared responsibility: the United Nations World Water Development Report 2, pp 243-274.

Jeffrey P, Gearey M (2006) Integrated water resources management: lost on the road from ambition to realisation? Water Science and Technology 53 (1): 1--8.

Mendoza GA, Martins H (2006) Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms. Forest Ecology and Management 230 (1-3): 1-22.

Mitchell B (2005) Integrated water resource management, institutional arrangements, and land-use planning. Environment and Planning A 37 (8): 1335--1352.

Mitchell VG (2006) Applying integrated urban water management concepts: A review of Australian experience. Environmental Management 37 (5): 589--605.

Niemczynowicz J (2000) Present challenges in water management: A need to see connections and interactions. Water International 25 (1): 139--147.

Pahl-Wostl C (2007) The implications of complexity for integrated resources management. Environmental Modelling & Software 22 (5): 561-569.

Pahl-Wostl C, Hare M (2004) Processes of social learning in integrated resources management. Journal of Community & Applied Social Psychology 14 (3): 193-206.

Postel SL (2000) Entering an era of water scarcity: The challenges ahead. Ecological Applications 10 (4): 941--948.

Stephens W, Hess T (1999) Systems approaches to water management research. Agricultural Water Management 40 (1): 3-13.

Sudmeier-Rieux K, Masundire H, Rizvi A, Rietbergen S (2006) Ecosystems, Livelihoods and Disasters: An integrated approach to disaster risk management. IUCN - The World Conservation Union, Cambridge, United Kingdom, 57 pp.

UNDP (2006) Human Development Report 2006. Palgrave Macmillan, New York, USA, 422 pp.

UNESCO-WWAP (2006) Water, a shared responsibility: the United Nations World Water Development Report 2. United Nations Educational, Scientific and Cultural Organization (UNESCO) & Berghahn Books, Paris, France, 584 pp.

Zehnder AJB, Yang H, Schertenleib R (2003) Water issues: the need for action at different levels. Aquatic Sciences 65 (1): 1--20.

# CHAPTER 2

## A SYSTEMS-BASED, GENERIC ENVIRONMENTAL DATA MODEL

# Preface

This chapter proposes a generic data model for environmental data. Within the context of the thesis, this data model bases on the theoretical principles developed in the SYSMOD systemic formalism (see Annex I), whose constructs are implemented into the data model. The resulting data model is tested with an implementation of the water management system model, defined in Chapter 1. It is a prerequisite for the subsequent software developments presented in the next chapters and in Annex II and III.

This chapter is adapted from the article: Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A (2010) A systems-based, generic environmental data model, submitted to Environmental Management.

# Abstract

Existing environmental data models have been designed to deal with limited ranges of data formats, or for the needs of specific applications. Therefore, when it comes to data integration, one has to deal with various data sources and different formats, which is a challenging and costly task. In this study, a generic data model for environmental data is proposed, which allows handling and integrating various environmental data. It is designed on the basis of a system modelling formalism, extended to support a wide range of information formats. To demonstrate its capabilities, the data model was applied to the two separate cases of (1) soil properties-related data and (2) electricity distribution-related data; the case of the water management system of Birmingham, United Kingdom, was furthermore fully implemented.

**Keywords:** data model, database, environmental data, data integration, integrated environmental management, systems approach, systems thinking

Data integration is necessary for enterprises and institutions wishing to gain access to centralised, transdisciplinary information. Achieving this integration is however a costly challenge (Bernstein and Haas 2008), as the data is produced by different services dealing with various tasks and is therefore often heterogeneous.

In the field of environment and natural resources management, geographical information systems (GIS) have become extremely popular over the last decades. As a result, most of the developments in data integration focussed on the creation and the dissemination of spatio-temporal databases (one can refer to the work of Pelekis (2004) to get an overview of the recent evolution in this field). Hence, the literature presents a profusion of such spatio-temporal data models. There are for instance data models specifically designed for supporting administrations in their management of natural resources (Crausaz and Musy 1997, Soutter and others 2009). Other models are rather designed as bases for environmental modelling (McKinney and Cai 2002, Cesur 2007, Strassberg and others 2007, Maidment 2008). Finally, some developments focus on the storage and retrieval of a wide range of interdisciplinary data as, for instance, the Observations Data Model (a relational model for environmental and water resources data focussing on numeric values dissemination (Horsburgh and others 2008)), the data model of the PANGAEA information system (Diepenbroek and others 2002), or the adaptive integrated data information system (AIDIS) which integrates remote sensing data (Flugel 2007). Notwithstanding the benefits of GIS, it is however obviously limited to a geographic objects-centred apprehension of the information and is therefore not appropriate for managing non-spatial data such as stakeholders-related information, legal references or interactions (such as fluxes or causal links), among others.

In order to integrate non-geographic elements and their related information, an alternative to the geo-centrist approach is a systemic organisation, composed of interlinked elements. Such systems-based data modelling approaches have been applied successfully to a wide range of specific domains (mainly related to networks) such as: water networks (McKinney and Cai 2002), social networks (Mitra 2007), dynamic transit networks (Huang and Peng 2008), engineering networks (Vishnyakov and others 2007) or graph-oriented databases (Erwig 1994, Kiesel and others 1995, Catarci 1996).

In summary, although certain fields of research have given rise to different approaches to data modelling based on systems, environmental data is generally modelled around geographic elements, and designed to deal only with specific kinds of data (such as numeric values) or with specific applications in mind (such as supporting hydraulic models). These data models are therefore not

adequate for a more generic integration purpose, gathering heterogeneous data including various formats and data not related to spatial elements. To fill this gap, this study proposes an innovative data model, relying on a core system structure adapted from a systems formalism. Environmental data from various domains can be organised into a system (as three case studies demonstrate) and can therefore be handled by the data model. Furthermore, the versatile system structure allows integrating elements from different domains. Hence, the proposed data model can be effectively used to manage and integrate various data from various domains, including the range of different formats such as multimedia files, fuzzy values, websites or opinions, and non-geographic data.

# Systems formalisms

As highlighted by Roquier and others (2010), existing visual formalisms for system modelling are limited when it comes to the description of interactions -such as fluxes or influences- between system elements. Roquier and others (2009) therefore developed an enriched formalism, called SYSMOD, to fill this gap. SYSMOD was considered very appropriate to represent environmental elements and data and was therefore adopted within this study.

SYSMOD defines system elements along two axes: the type and the level. The type can be either “node”, “interaction” or “information”. The level can be either “class”, “instance” or “property”. A node is a system element, such as a stakeholder, or a river. An interaction is a functional relation between two groups, such as a flux or an influence. And an information is an actual value such as a numeric data or a text.

Regarding levels, a class is an abstract template providing a definition for any instance belonging to it, and conversely an instance is an actual element whose category is defined by its class. For instance, a given, real water treatment works is an instance of the class water treatment plant. Finally, properties are references to other elements. For instance, the previously mentioned water treatment works may “have” one or more sand filters, and it would therefore have a property called “sand filters” referencing one or more instances of actual sand filters. A property that references nodes or information acts as a list of instances, and is therefore called a “group”. This covers nodes

groups and information groups. An interaction can also be a property but cannot be considered as a group or an instance (it can be a class, though).

Overall, combining the two axes (type and level) gives rise to eight different constructs (interaction being not represented as an instance). Furthermore, five types of structural relations are defined, as follows:

- Has property: This is a relation between any element and a property.
- Specialisation: This is the relation between a class and a subclass. For instance, a black soil (chernozem) is a subclass of the very general soil class. It therefore allows defining taxonomies.
- Instantiation: This is the relation between a class and one of its instances.
- Property instantiation: This is the relation between a class and one of the properties deriving from it.
- Group member: This is the relation between a group and a member of a group (an instance).

The visual representation of all these constructs is given in Table 1.

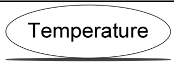
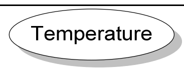
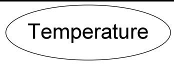
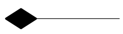
| Definition | Representation |
|---|---|
| Node class |   |
| Node group |   |
| Node (instance) |   |
| Information class |   |
| Information group |   |
| Information (instance) |   |
| Interaction class |   |
| Interaction (group or instance) |  |
| Has property |  |
| Specialisation |  |
| Instantiation |  |
| Property instantiation |  |
| Group member |  |

Table 1: SYSMOD visual elements

# Model

### Design

Many conceptual frameworks are available for designing data models, providing general methodologies and more specific tools. The framework used for this study is Perceptory, developed by Laval University (Bedard and others 2004, Bedard 2005). Based on the versatile unified modelling language (UML, which was originally created for software design), Perceptory enriches the latter with semantics specific to data modelling. It is a free plug-in for Microsoft Visio diagramming software, which includes support for modelling the space and time dimensions. It also provides a code generator allowing the implementation of the data model into a database schema. The UML-based formalism used by Perceptory for representing entities and relationships is explained in Table 2. Spatio-temporal aspects are modelled visually using pictograms as shown in table 3. Although this formalism show similarities with SYSMOD (Table 1), its aim is the design of databases, and the underlying UML is a widely accepted -and therefore easy to read- language by the people working in this field.

### Overview

The core of the data model is presented in figure 1; figure 2 contains the domains, namely enumerations of the values that an attribute can take; altogether, this includes five main packages:

1. The system package contains the core of the model: system elements, based on the SYSMOD formalism.
2. The information package contains different types of information: value types and values, texts, document files, geometries and lifetimes. Information is in fact a particular case of system elements.
3. The themes package provides a way of creating groups of system elements such as indicators sets or scenarios.
4. The data processing package provides a way of documenting the origin of data, either being created through expertise or modelling.
5. The domain package contains all the domains used throughout the data model.

The next sections provide details about the system and information packages. They are followed by a complementary important topic: the inclusion of uncertainties in the data model.

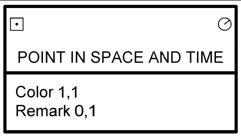
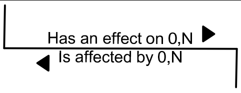
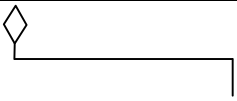
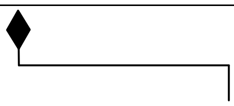
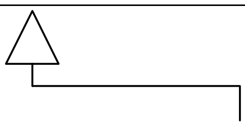
| Diagram Element | Description |
|---|---|
|   | Entity: This is implemented as a table in the database. With perceptory, the geometry and temporality can be defined using pictograms (on the top-left and top-right of the entity, see also table 3), in this case a point in space and time. Attributes define the fields of the database table. The values beside them indicate the minimum and maximum numbers of instances the attribute may have. A zero indicates that the attribute is nullable. |
|   | Association: This provides a functional relationship between entities, implemented as sets of primary keys and foreign keys, possibly involving an association table (in many to many cardinality relationships). The cardinality of the relationship is indicated by two numbers at each end of the association. It defines how many linked entities may be involved in the association. |
|  | Aggregation: This is a special case of association. It represents a “has” functional relationship where one entity on the white diamond end “possesses” many (one to many cardinality) entities on the other end. Destroying the possessor involves the destruction of the possessees. |
|  | Composition: This is also a special case of association. It represents an “is” functional relationship, where one entity on the black diamond end is “composed” by many (one to many cardinality) entities on the other end. Destroying the composer involves the destruction of the composees. |
|  | Generalisation: This is a hierarchical relationship between one parent and many children (one to many cardinality). The children “inherit” attributes and relationships their parents have. There are different ways of implementing generalities in databases, including the implementation of both parent and children as separate tables, or the creation of a unique table gathering parent and child. |

Table 2: The formalism of Perceptory: an extension of the unified modelling language (UML), used for designing databases (while SYSMOD is used to describe systems)

| Pictogram | Meaning |
|---|---|
|  | Spatial : any geometry |
|  | Timestamp: duration |

Table 3: Partial list of the spatio-temporal pictograms of Perceptory

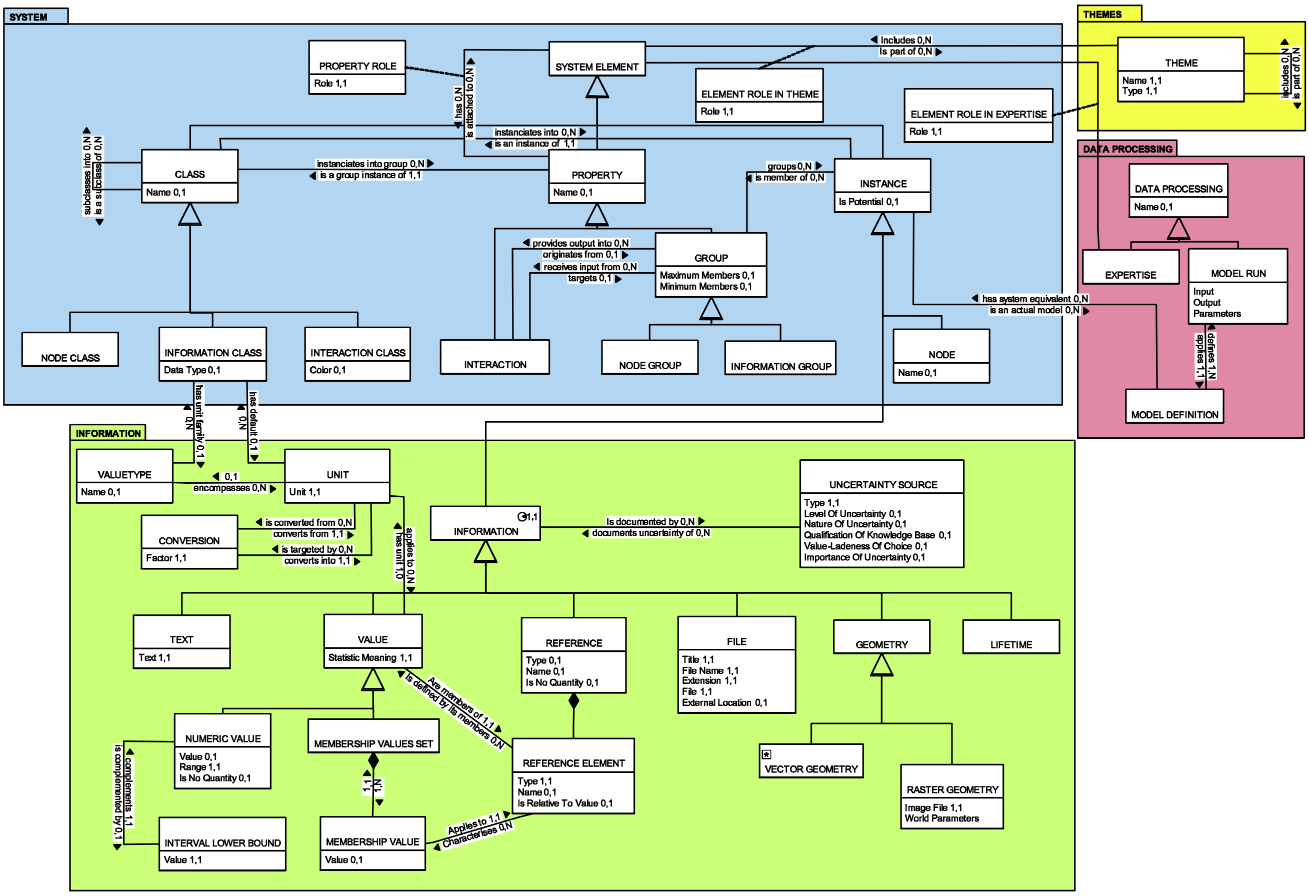



Figure 1: Core data model

## System components

The system package features the eight constructs defined within the SYSMOD formalism. These eight elements are all specializations of the “System Element”, which is then subdivided along the three different levels (class, property and instance). The “Group” entity is a further subdivision of “Property”, excluding “Interaction”. Indeed, the latter is defined as a relation between two groups. Regarding the attributes of these entities, all the constructs have a name (except information). Groups are documented by the minimal and maximal number of their members. Instances are defined as potential (for non-real elements, used in scenarios for the future) or normal. Finally, the information class has a data type attribute defining if the associated data is to be stored as a file, a value, a geometry, etc. (this domain covers the different kind of information in the information

package). Furthermore, information classes are associated to their possible units (through "Value Type") and to a default unit (relation to "Unit" entity).

Beside system elements, the system package features the five different structural relations defined in SYSMOD. The "Has property" relation is modelled as the associative entity "Property Role" linking any system element with a property. The attribute of this entity enables specifying the role of the property ("Property Roles Domain" values such as: "Has", "Manages" or "Uses"). The self-relation on the "Class" entity allows linking a base class to a subclass (specialisation relation). The instantiation and property instantiation structural relations are modelled by the link between "Class" and "Instance" or "Property" respectively. Finally, members of a group are gathered into their group through the "Instance" to "Property" relation.

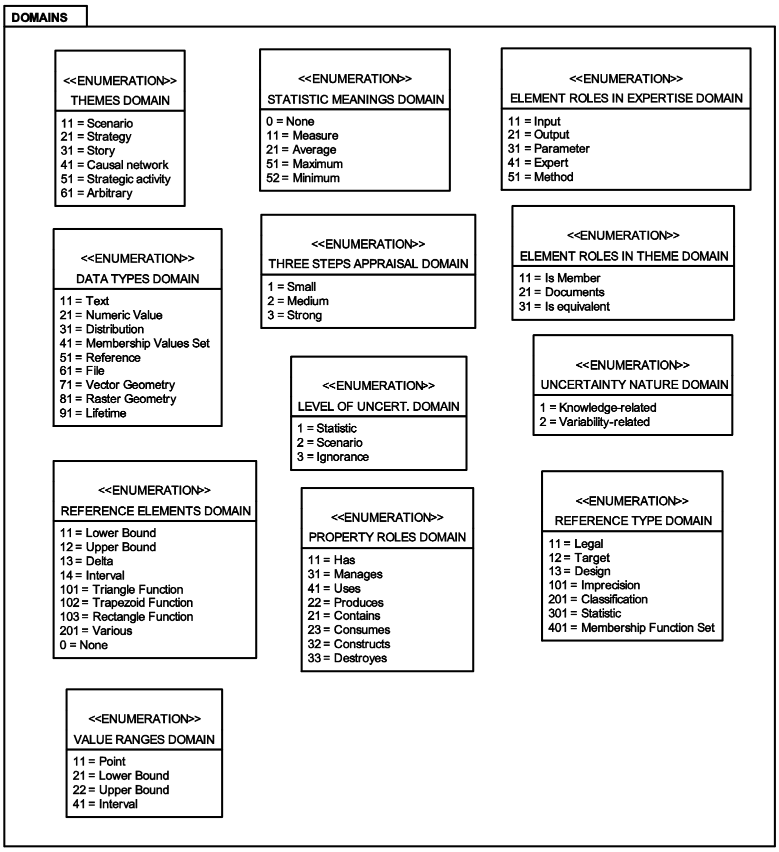



Figure 2: Data model domains

### Information

The information package covers a large range of different kinds of quantitative and qualitative data types, under the generic "Information" entity. The latter is itself a system element, as it derives from "Instance". But unlike the other system elements, it is further refined into several specialized entities reflecting the various categories of information. Any information data is an instance of an information class which defines its type through its name (such as "e-mail address", "DBO5 concentration", "vector geometry", etc.), its possible units ("Value Type" and "Unit") and a data type. The data type reflects the six categories of information, as described hereafter:

- Text: Many different kinds of information can be stored as texts, such as addresses, comments, descriptions, etc. The "Text" entity simply allows storing such a text in its single attribute.
- Value: The data model supports values in two different formats: numeric and fuzzy values. Numeric values can either be a crisp value, a lower or upper bound, or an interval, as defined by the "Range" attribute domain value. In the interval case, the complementary value of the interval is provided by the auxiliary "Interval Lower Bound" entity. Fuzzy values are defined by a set of membership values ("Membership Values Set", composed of "Membership Values"). The latter apply to membership functions defined in the "Reference" entity, and whose "Type" attribute takes the value "Trapezoid", "Triangle" or "Rectangle".

  The "Statistic Meaning" attribute allows specifying if the value is a basic observation, an average, a maximum, etc.

  Furthermore, it might sometimes be useful to mention values, whose actual quantification are unknown, by leaving the "Value" field null. For instance, there might be missing values in an air quality survey at different locations in a city. Inserting a new empty value emphasizes that the actual result is missing.
- Reference: There are different kinds of references. They can document limits (legal, target or design), imprecision, classification, further statistic features (such as covariance or regression values, which don't directly characterize the imprecision), or they can be membership functions. All these different kinds of references require characterisation through values. To define limits, only one value is required (upper or lower bound). Imprecision may be defined by one or more values, through a delta (plus or minus), an interval, percentiles, or by non-structured, loose values, such as its variance. Classification requires the definition of intervals, corresponding to some interpretations. Membership functions need the

construction of functions (trapezoid, triangle, etc.) through a few constitutive parameters. The values required to define such references are stored as any other value in the "Value" entity. They endorse their role of reference through the relation to the "Reference Element" entity, which specifies their role. For instance, a temperature classification may include four ranges: very cold, cold, hot and very hot; the temperature values 0°C, 15°C and 30°C playing the role of interval bounds (or parameters of the membership functions if fuzzy).

In some cases, references are not based on actual physical values. For instance, temperature appreciations might be defined qualitatively by surveyed people, with an arbitrary scale ranging from "tropical" to "freezing", without any definition of the actual bounds. In this case, the classification marks might just be a scale from, say, 1 to 10. To distinguish physically based, quantitative references from qualitative ones, the "Is No Quantity" attributes of "Reference" and of "Numeric Value" are available.

- File: Files can be stored in binary format (binary large objects – BLOB), directly in databases. This allows handling any kind of files, including multimedia items: videos, audio, pictures, documents, etc. If the file is not to be stored directly in the database (e.g. it is too big or it is not in digital format), its location can be referenced using the "External Location" optional attribute.
- Geometry: Geometries can be either features (vector elements) or coverages (raster data). In the former case, the geometry can be stored as a text (e.g. well-known text) or as binary. In the latter case, the coverage is defined by the combination of the image itself and possibly by complementary parameters specifying its spatial location and extent.
- Lifetime: This information simply documents a "birth" and "death", a validity period.

### Uncertainties

In the data model, uncertainties are documented through uncertainties background data, and through imprecision values. The former, uncertainty background data, is stored in the "Uncertainty Source" entity. It is based on the concept of "uncertainty matrix" defined by Walker (2003) and adapted by Janssen (2005), a recognized approach which has been used in information management projects such as HarmoniCA (Refsgaard and others 2007). An uncertainty matrix defines six dimensions: (i) the source of uncertainty (such as the base data or the model specificities -defined in the "Type" attribute), (ii) the level of uncertainties (statistical, scenario, ignorance), (iii) the nature of uncertainty (related to knowledge or variability), (iv) the qualification of the knowledge base (from small to strong), (v) the value-ladenness of the choice (qualifying factors such as the emotional

background behind the choice –from small to strong), and (vi) the importance of the uncertainty (as perceived by the user).

Furthermore, the uncertainty background may also be documented using concepts such as pedigree matrices (van der Sluijs and others 2005). These matrices are groups of evaluations against given criteria such as the consensus between involved experts, the quality of the input data in the model, or the quality of the model used. Each evaluation may be stored using the "Reference" and "Reference Element" mechanism (described in the information-related previous point). And matrices may be formed by grouping the evaluations into themes (as "Indicators set" in the "Type" attribute of a "Theme").

The second kind of uncertainties supported by the data model, imprecision, may be documented through a "Reference" having a value "Imprecision" in its "Type" attribute. The reference may then encompass any number of percentiles, intervals, deltas, or bounds.

# Application

To demonstrate the usability of the data model, the case study of the water management in Birmingham, United Kingdom, was fully implemented. Furthermore, to demonstrate the genericity of the model for various environmental data, two additional cases were studied: soil properties and electricity distribution-related data.

### Soil-related data

Figure 3 shows a systemic model of soil-related data, as well as some derived elements, linked to the fictitious Green City. In the upper part, the classes of the systemic model focus on the definition of a soil, documented by a type, vegetation and one or more soil layers. The latter show numerous related information groups such as depth, texture, colour, etc. Plot, equipment and vegetation are rather satellite elements and are therefore not extensively documented. In the middle of Figure 3, the base soil classes are refined (specialised) into the definition of a specific category of soil, the chernozem. Chernozems are a kind of soil (a subclass of soil). But more specifically, they always have

high organic matter content and therefore appear quite black. Finally, some fictitious instances of soil, plot and vegetation are shown in the lower part of Figure 1. They don't repeat the information groups properties of their respective classes; this would be similar.

This small case study demonstrates that soil data can be successfully organised into a systemic model, using the system elements constructs implemented into the data model.

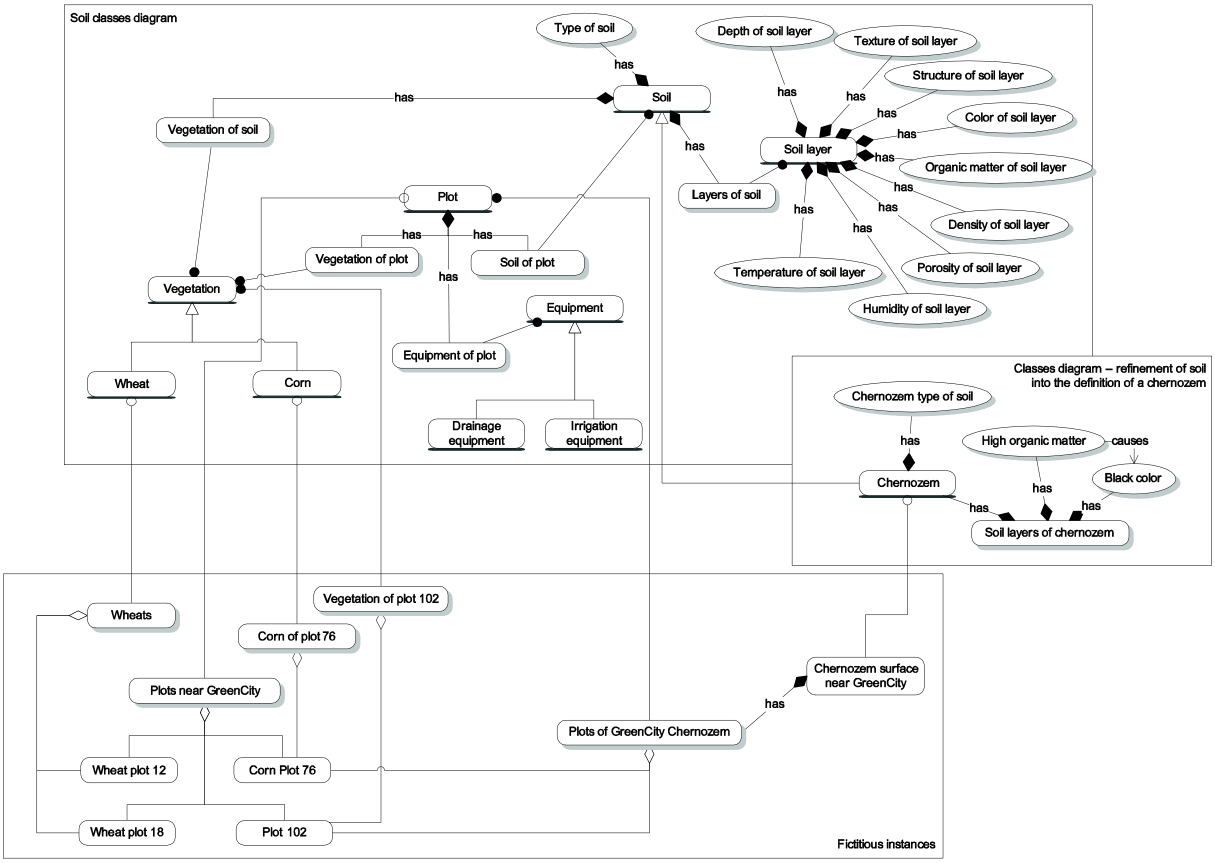



Figure 3: Soil properties systemic model

### Electricity distribution data

Figure 4a shows a very simple systemic model for the distribution of energy at a very high level, without many details: the energy resource is transformed into electricity by a power plant; the transmission network brings the electricity to transformers which provide low voltage current to final consumers. Besides, in the top-most part of Figure 4, the electricity interaction class is defined; it is documented by two information groups: voltage and the actual electricity flux.

The rough model in Figure 4a is refined into a more detailed one shown in Figure 4b. This refined model shows that the distribution network may be decomposed into different parts with different voltages, and different kinds of consumers. It also adds details about some kinds of power plants, which may feed the transmission grid at different voltages.

Again, this small case study can be modelled using the system constructs and its related data can therefore be handled by the proposed data model. Furthermore, it illustrates that different levels of details of a same topic can be handled by the data model.

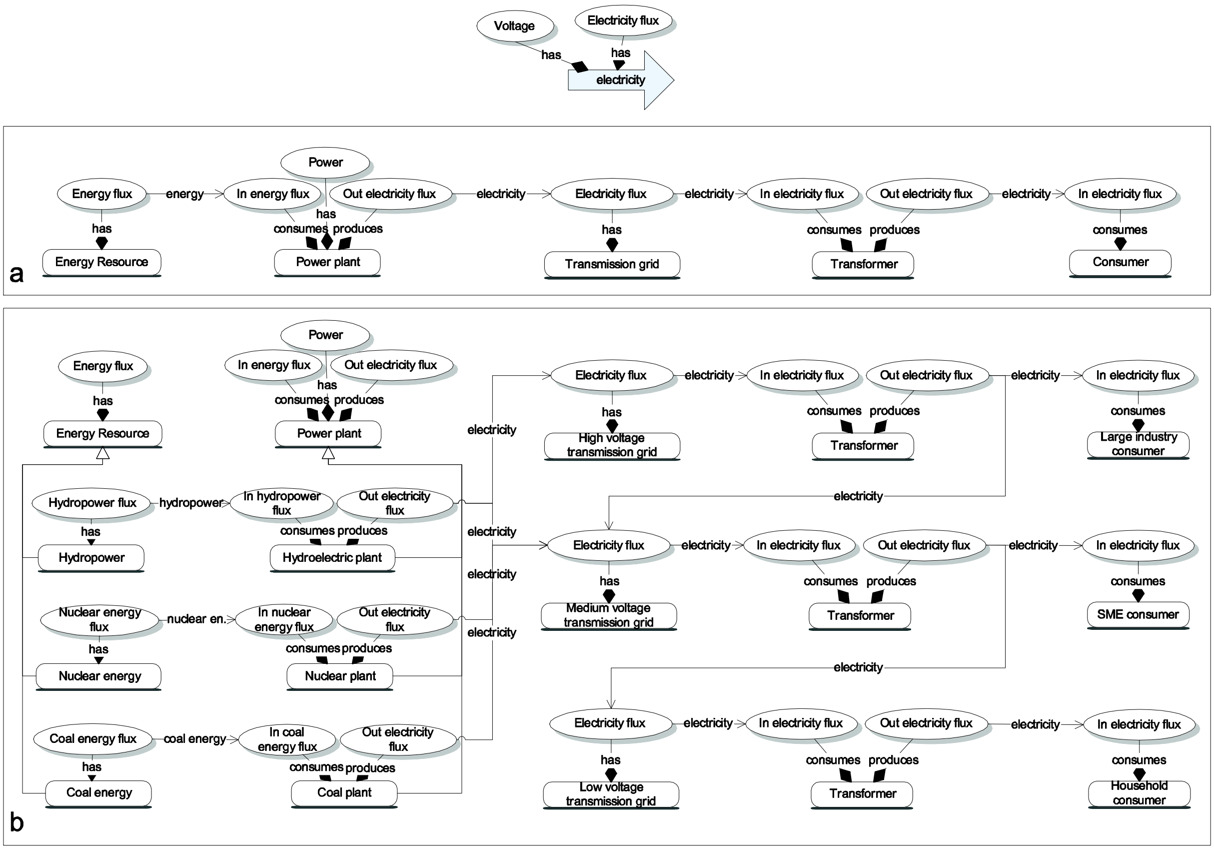



Figure 4: Energy distribution systemic model

## Water management in Birmingham

For a more extensive test, the data model was fully implemented as a database and fed with data related to the water management system of the city of Birmingham in the United Kingdom, accounting almost 5 million inhabitants.

The data model was implemented in the open-source database management system (DBMS) PostgreSQL (Group 2009), with the PostGIS (Community 2009) extension to manage the geographic data. This was achieved using the code generation functionalities in Perceptory, which allow creating an Oracle DBMS schema. The Oracle schema was then converted to PostgreSQL.

To define a systemic model, the water management system model proposed by Schenk and others (2009) was used. This model organises graphically water management-related components (including elements beyond the strict field of water, such as stakeholders, lakes, funds or laws) and their interactions into a system. This model was easily converted into a SYSMOD formalism-based systemic model, which was then introduced into the database. This led to the creation of almost one hundred node classes organised into a tree connected by “specialisation” relations. Several interaction classes were also extracted out of the model: influences (functional relations defined by a text) and different fluxes (water, money, waste, data, material fluxes). Finally, numerous information classes were created on the basis of usual water-related indicators sets and organised into a tree of indicators.

The database was then fed and consulted using a prototype interface allowing focusing on a system element, documenting it and editing its relations. This way, about two hundreds elements were added including nodes or node groups (e.g. stakeholders, water works or laws), interactions (e.g. fluxes and influences), information and information groups (e.g. flux values, web sites, files, chemical concentrations, geometries - in well-known-text WKT format). The elements were identified through an analysis of the water management system of the city.

The application of the data model to the water management system of Birmingham demonstrated that it can be implemented into a functional database, adequately hosting various kinds of data from different fields (such as laws, stakeholder data, pictures, various geometries, text files, etc.).

# Discussion

### Other approaches

Data integration can be achieved at different architectural levels, spanning from data storage centralization to a “manual” integration by the end-users, with intermediate possibilities within

applications: in the data access layer, the middleware, the client code, or in the user interface. (Ziegler and Dittrich 2004, Halevy and others 2005, Bernstein and Haas 2008). At the level targeted by this study, -the lowest, data storage level-, database management systems (DBMS) provide technologies enabling highly-structured data storage, enforcing rules and constraints (in comparison, in other storage systems such as spreadsheets or file systems, the data structure is looser). There are different database models available, including the relational, the object and the multidimensional models. The first one is a well-established technology, with numerous related products available. However, it presents a mismatch (known as the impedance mismatch) with programming languages, which are generally structured into objects (object-oriented code). To address this issue, the object database model proposes an object-oriented structure. Albeit this is an interesting technology, it is more recent, and much less spread (Leone and Chen 2007). Finally, in the multidimensional model, the data is organized in multidimensional cubes, which are particularly well-suited for analysis (Pedersen and Jensen 2001), for instance with OLAP (online analytical processing) technologies. In the present study, the relational model was selected, as it was considered more appropriate to the highly relational structure of the developed data model than the multidimensional structure, and more widely accepted than the object model.

Database management systems enforce a strict structure through a schema, which includes rules and constraints. Ontologies propose an alternative approach. In ontologies, there is no underlying schema. The data is meta-documented using a set of descriptive tags which also allow defining inter-relations. This way, any kind of data can be added, linked to existing data, and thereby integrated. An example of this approach is the semantic web project, in which heterogeneous and distributed data available on the internet is to be meta-documented using the OWL formalism (Web Ontology Language). Then, specific applications shall be able to interpret these meta-documentations and thereby extract and provide relevant information as requested by the users. Hence, ontologies provide an extensible, open-world framework, which allows partial or even whole inconsistencies in the data (Horrocks 2008).

In the present study, the schema-based approach was preferred to ontologies for a few reasons: (i) ontologies are open-world, which makes them less appropriate to enforce a strict structure (a schema), and therefore to enforce the system formalism defined by SYSMOD, (ii) ontologies are relatively recent and there are only a few, prototypic tools available to manage them, (iii) databases perform better when it comes to answer queries, because ontologies can grow very complex and necessitate logical analysis to explore their open-world structure.

Database schemas, on the other hand, are not versatile: modifying them may involve a subsequent reorganisation of the data, as well as an adaptation of the possible interfaces performing requests in the database. The hereby adopted application of system modelling constructs into database entities allows to partially overcome this limitation, because the system structure is generic and because it is further completed by the inclusion of the various formats of information: texts, values, files, geometries, etc. This generic structure fits to various environmental data (organised as systems), which can therefore be handled without the need to modify it.

### Achieving integration

The inclusion of system modelling constructs into the data model does not only allow handling various environmental data. It also allows handling different underlying perceptions -different systemic models- over different fields and their related data. This is illustrated with the electricity distribution case study featuring two different perceptions of the topic. In a database implemented on the basis of the proposed data model, such complementary models can coexist. And concurrent models can also coexist, as well as various semantics. For instance, a dataset could have both a node class called “Transmission grid” as well as a node class called “Transmission network”. To avoid redundancy in the actual information (numeric values, files, etc.) both nodes could be connected to the same information groups.

Hence, the proposed data model allows handling various kinds of data and different perceptions from different fields. But moreover, these pieces of data can be linked, integrated. For instance, in the electricity distribution case, a hydropower plant is defined. Such an element also exists in the water management system model case. Both elements could be merged. Or if they are conceptually or semantically different, they could be redefined as two different subclasses of a common element. Ultimately, links can be created this way between neighbouring fields and their related data, leading to the actual integration of the data into a global system.

### Potential users

Although the proposed data model has the ability to handle various environmental data and various perceptions, implementing it would not be systematically recommended. For specific applications, targeting a well-defined need, with potentially no or little changes over time in the data structure, a dedicated data model would be simpler, closer to the actual need, and would therefore perform better. In such cases, the proposed data model would probably prove a too complex and heavy

structure. This typically encompasses small businesses or boards in charge of a specific mission with no need to export its data towards multiple external users. The proposed data model would rather be recommended for larger organisations, where integrating the data is for more general, open purposes. This includes at least large environmental boards with diverse sub-sections managing various data, or planning offices gathering heterogeneous strategic data.

# Conclusion and perspectives

Models developed so far to manage environmental data were mostly designed for specific tasks, and none was meant to deal with a large range of various information formats, with non spatial-centric objects, and with concurrent (or complementary) perceptions of a domain specific data. The data model proposed in this study innovatively addresses these issues through an underlying core system structure extended to handle various data formats, making it generically suitable for the various environmental domains' data. Texts, numeric values (including fuzzy), references, files (e.g. documents, multimedia, audio), geometries and dates are included. Elements related or not to a spatial object are managed similarly -as system elements; this makes it easy to include data such as stakeholder's names and addresses, money fluxes, legal documents or causal links. Different systemic models of the same field can coexist and can ultimately reference the same information (values, texts, etc.); this makes it possible to provide different perceptions upon a field to various end-users including also applications such as calculation models (in a given field, available calculation models often have different input and output data structures). Finally, links between different systemic models for different fields can be defined and created, thereby merging the models, the datasets, and achieving transdisciplinary integration.

For enterprises and institutions active in environmental management and wishing to integrate data from various sources, the proposed data model enables the deployment of a unique, centralised database generically able to handle the data of the different departments. Such a database doesn't need to be modified whenever data structures change or new ones are added, which dramatically eases the maintenance needs of the database itself, and also of the applications querying it; only the

underlying systemic model would evolve (which also requires some -but far less- maintenance, at a more at a conceptual level).

Deploying databases integrating various environmental data enables the provision of holistic views and thereby a global apprehension over complex environmental problems. In that regard, further developments shall focus on developing methodologies and technologies to harness the datasets stored in such integrated databases, in order to be able to conveniently retrieve, select and display the information. This could be achieved using graphical user interfaces displaying the systems-based information to the user as a network of connected nodes and relations. Another interesting perspective, enabled by the genericity of the data model, is the development of an interface allowing interoperability with various models, which could thereby exchange inputs and outputs on a common basis.

# References


Bedard Y (2005) Perceptory 2006, a conceptual modeling tool for geospatial databases. Accessed online 01 / 30 / 2008: http://sirs.scg.ulaval.ca/perceptory/english/enewindex2.asp.

Bedard Y, Larrivee S, Proulx MJ, Nadeau M (2004) Modeling geospatial databases with plug-ins for visual languages: A pragmatic approach and the impacts of 16 years of research and experimentations on perceptory. In: (ed) Conceptual Modeling for Advanced Application Domains, pp 17-30.

Bernstein PA, Haas LM (2008) Information integration in the enterprise. Communications of the Acm 51 (9): 72-79.

Catarci T (1996) A graph based framework for multiparadigmatic visual access to databases. IEEE transactions on knowledge and data engineering 8 (3): 455-475.

Cesur D (2007) GIS as an information technology framework for water modeling. Journal of Hydroinformatics 9 (2): 123-134.

Community P (2009) PostGIS. Accessed online 01 / 09 / 2009: http://postgis.refractions.net/.

Crausaz PA, Musy A (1997) GESREAU, an institutional GIS for integrated water management. In: Baumgartner MF, A. SG and I. JA (ed) Proceeding of International Symposium on Remote Sensing and Geographic Information Systems for Design and Operation of Water Resources Systems, at 5th Scientific Assembly of IAHS, Rabat, Morocco, pp 33-41.

Diepenbroek M, Grobe H, Reinke M, Schindler U, Schlitzer R, Sieger R, Wefer G (2002) PANGAEA - an information system for environmental sciences. Computers & Geosciences 28 (10): 1201-1210.

Erwig M (1994) Explicit graphs in a functional-model for spatial databases. IEEE transactions on knowledge and data engineering 6 (5): 787-804.

Flugel WA (2007) The adaptive integrated data information system (AIDIS) for global water research. Water Resources Management 21 (1): 199-210.

Group PGD (2009) PostgreSQL. Accessed online 01/09/2009: http://www.postgresql.org/.

Halevy AY, Ashish N, Bitton D, Carey M, Draper D, Pollock J, Rosenthal A, Sikka V (2005) Enterprise information integration: successes, challenges and controversies. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, ACM: Baltimore, Maryland.

Horrocks I (2008) Ontologies and the Semantic Web. Communications of the Acm 51 (12): 58-67.

Horsburgh JS, Tarboton DG, Maidment DR, Zaslavsky I (2008) A relational model for environmental and water resources data. Water Resources Research 44 (5): 12.

Huang R, Peng ZR (2008) A spatiotemporal data model for dynamic transit networks. International Journal of Geographical Information Science 22 (5): 527-545.

Janssen P (2005) A guidance for assessing and communicating uncertainties. Water science and technology 52 (6): 125-131.

Kiesel N, Schuerr A, Westfechtel B (1995) Gras, a graph-oriented (software) engineering database system. Information Systems 20 (1): 21-51.

Leone A, Chen DY (2007) Implementation of an object oriented data model in an information system for water catchment management: Java JDO and Db4o Object Database. Environmental Modelling & Software 22 (12): 1805-1810.

Maidment DR (2008) Bringing water data together. Journal of Water Resources Planning and Management-Asce 134 (2): 95-96.

McKinney DC, Cai XM (2002) Linking GIS and water resources management models: an object-oriented method. Environmental Modelling & Software 17 (5): 413-425.

Mitra S (2007) Design of a data model for social network applications. Journal of database management 18 (4): 51-79.

Pedersen TB, Jensen CS (2001) Multidimensional database technology. Computer 34 (12): 40-+.

Pelekis N, Pelekis (2004) Literature review of spatio-temporal database models. The knowledge engineering review 19 (3): 235-274.

Refsgaard JC, van der Sluijs JP, Hojberg AL, Vanrolleghem PA (2007) Uncertainty in the environmental modelling process - A framework and guidance. Environmental Modelling & Software 22 (11): 1543-1556.

Roquier B, Schenk C, Soutter M, Mermoud A (2009) A system modelling formalism for environmental information. Article under preparation.

Roquier B, Schenk C, Soutter M, Mermoud A (2010) SYSMOD: A systems modelling language for environmental information. Submitted to Ecological Modelling.

Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management. Environmental Management 43 (3): 458-469.

Soutter M, Alexandrescu M, Schenk C, Drobot R (2009) Adapting a geographical information system-based water resource management to the needs of the Romanian water authorities. Environmental Science and Pollution Research 16 (S1): 33-41.

Strassberg G, Maidment DR, Jones NL (2007) A geographic data model for representing ground water systems. Ground Water 45 (4): 515-518.

van der Sluijs JP, Craye M, Funtowicz S, Kloprogge P, Ravetz J, Risbey J (2005) Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System. Risk Analysis 25 (2): 481-492.

Vishnyakov VY, Gritsenko YB, Zhukovsky OI (2007) The data model for the topology relations representation in the engineering networks. In: (ed) Proceeding of IEEE International Siberian Conference on Control and Communications, Tomsk, RUSSIA, pp 89-93.

Walker WE, Harremoes P, Rotmans J, van der Sluijs JP, van Asselt MBA, Janssen P, von Krauss MPK (2003) Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment 4 (1): 5 - 17.

Ziegler P, Dittrich KR (2004) Three decades of data integration all problems solved? In: Jacquart R (ed) Proceeding of 18th World Computer Congress, Toulouse, FRANCE, pp 3-12.

# CHAPTER 3

## A SYSTEMS-BASED INFORMATION SYSTEM

# Preface

This chapter proposes an information system (IS) based on a systemic formalism. Within the context of the thesis, this IS applies the data model defined in Chapter 2, which itself leans on the SYSMOD systemic formalism (in Annex I). The information system is tested with an implementation of the water management system model, defined in Chapter 1, enriched graphically by icons, specifically designed for each inventoried system element. It is a prerequisite for the subsequent software developments presented in the last chapter and in Annex II and III.

This chapter is adapted from the article: Schenk C, Roquier B, Soutter M, Mermoud A (2010) A systems-based information system, submitted to Information Systems.

# Abstract

Environmental management is lacking information systems (IS) able to generically handle various data from different fields in various formats. For instance, geographic information systems (GIS) propose a map-based, geo-centrist apprehension of the information and are therefore not well suited to manage non-spatial data (such as laws or stakeholders) and interconnections. Organising data in a more fundamental way, along a system structure, including nodes and interconnections, allows overcoming such limitations. Although there are graphical tools to describe systems, they do not offer the possibility to associate system elements with data stored in a database. This study proposes an innovative systems-based information system (SIS) meant to bridge this gap between system description and information systems. The SIS includes a database that has itself a system structure which allows handling any data modelled as a system (in a systemic way). Therefore, a first benefit of the SIS is the ability to manage data generically through a user-friendly interface. Secondly, it provides a framework which fosters fathoming and communicating perceptions and data of complex situations through the creation, the documentation and sharing of thematic system views.

**Keywords:** environmental data, systems-based information system, systems thinking, data integration, integrated management

# Introduction

An information system can be defined as an integrated set of components for collecting, storing, processing, and communicating information. In practice, this covers a host a various applications in different fields. In the field of environmental management, the prominent form of information systems is probably the geographic information system (GIS) which proposes a map-based user interface. Indeed, the past years have seen dozens of publications describing environmental GIS. Most recent developments include water catchment GIS (Nishihama and others 2007, Soutter and others 2009), soil and agriculture-related GIS (Klass and others 2007, Li and others 2007, Pradhan and others 2008, Jorge 2009) or benthic environment GIS (Shih and others 2009). Another popular, rawer form of information system is the data warehouse, which doesn't necessarily include a graphical user interface (whereas GIS present a rich user interface, generally backed by a kind of database). There is also a host of publications relating to such developments. For example, in the last years, data warehouses were deployed in the fields of agriculture (Sander and others 2009), aquatic ecosystems (Peredo-Parada and others 2009), hydrogeology (Comunian and Renard 2009), foods' properties (Kavvadias and others 2009) and pharmaceutical products in the environment (Cooper and others 2008, Miege and others 2009), to mention just a few.

Albeit GIS are powerful to display spatial elements, they are more limited when it comes to non-spatial data, although the latter can be linked as attributes of the spatial elements. But this implies a way of organising the data around spatial elements, a geo-centrist data structure. Furthermore, GIS are not well suited to represent interconnections (fluxes, causal relations), which makes it difficult to display for example a city's water cycle with its interlinked elements (e.g. treatment works, consumers, wastewater collection). Data warehouses are not constrained by a geo-centrist user interface, but they are always geared towards specific needs and applications and therefore lack genericity (Schenk and others 2010).

A generic way of organising the data in a database is through a systems-based structure, because such a structure is able to host any systemic model, and because any environmental data can be modelled into a systemic model (Schenk and others 2010). Such systemic models can be created graphically using a visual formalism in which elements and interconnections between elements can be defined. As this systems approach allows defining also non-spatial elements and interconnections,

it provides a more fundamental framework than the geo-centrist data structures of GIS. However, existing system modelling tools are not designed to deal with data and databases; they are designed only to draw diagrams.

To bridge the gap between information systems and systemic models, this study proposes an innovative systems-based information system (SIS) which uses the expressive power of a systemic visual formalism representation backed with a database. The latter is itself an implementation of a systems-based data model, generic for environmental data. Therefore, the benefits of the SIS are twofold. Firstly, it provides a graphical user interface (GUI) which allows managing a generic database for environmental data. This means that various data (such as energy- or soil- or water-related data) in various formats (such as numeric values, texts and multimedia files) can be integrated in a user-friendly way and into a single database. Secondly, it allows creating and visualising systemic models and therefore acts as a platform for formalising, sharing and communicating perceptions of possibly complex situations, through their representations as system diagrams with their related data.

# Systems-based data model

The hereby proposed systems-based information system leans on the previous development of a systems-based data model (Schenk and others 2010) which in turn implements a system formalism called SYSMOD (Roquier and others 2010).

In SYSMOD, system elements are defined along two axes: the type and the level. The possible types are “node”, “interaction” and “information”. A node is a system element, such as a lake, a person or a policy. An interaction is a flux or an influence between two groups. Finally, an information is an actual value whose format can be a numeric value, a text, a geometry, a file, or a lifetime. The possible levels are “class”, “instance” and “property”. A class is an abstract template allowing the creation of instances based on its definition. For example, a given, real-world lake is an instance of the “lake” class defining what a lake is. Properties of an element are classes used within the context of an element. For example, a car may have the properties “number of wheels”, a “colour”, etc. The latter are instances of the wheel and colour classes, within the context of a car. As the cars have

more than one wheel, and maybe different colours, these properties are also called "groups". Interactions are also properties, but cannot be considered as a group or an instance (it can be a class, though). Overall, the combination of the two axes (type and level) produces eight different constructs (interaction being not represented as an instance). Furthermore, five types of structural relations are defined: has property (the relation between any element and a property), specialisation (the relation between a class and a subclass – for example a teacher could be the subclass of a person), instantiation (the relation between a class and one of its instances), property instantiation (the relation between a class and one of the properties deriving from it), group member (the relation between a group and a member of a group).

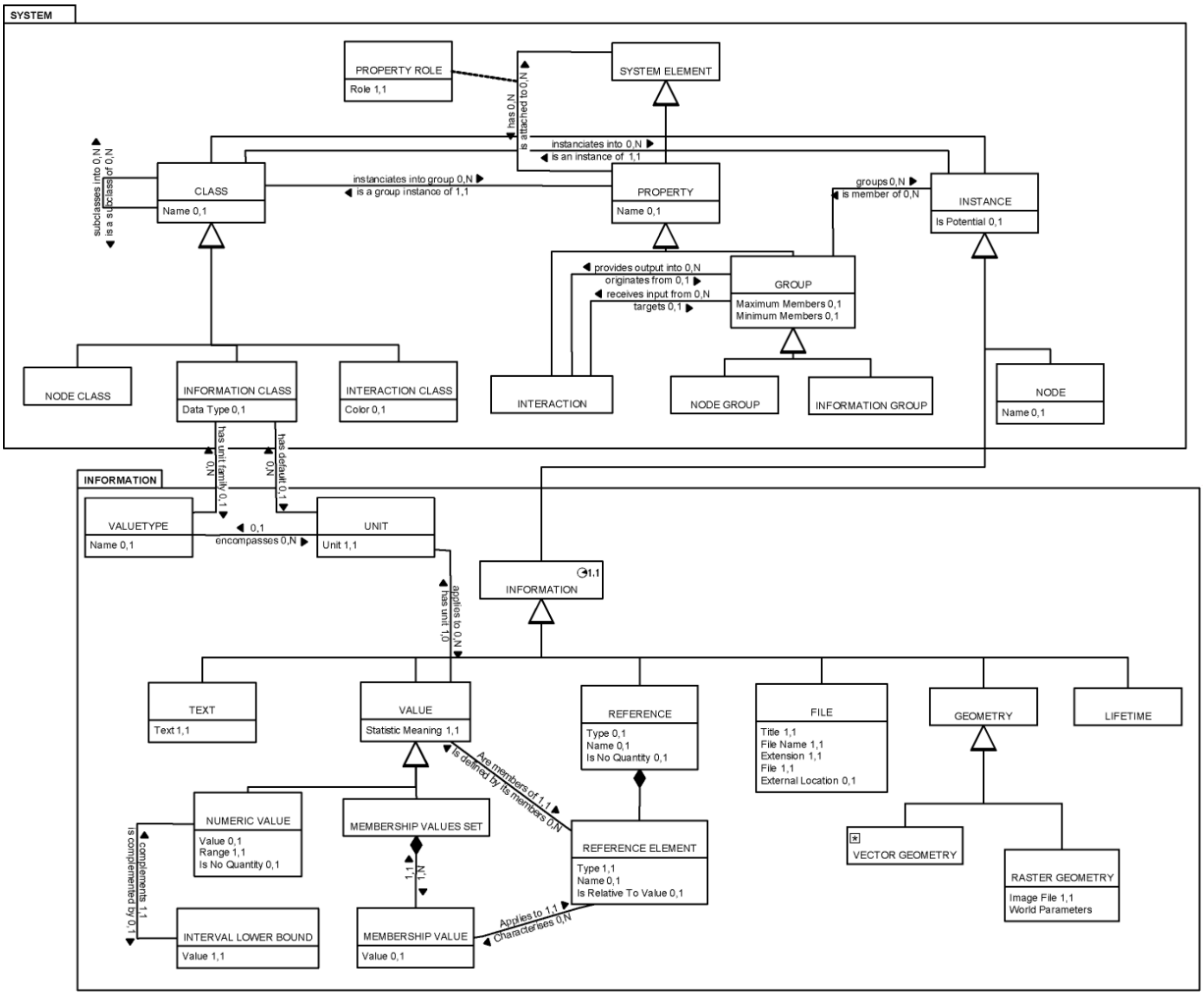



Figure 1: Database's data model core structure (adapted from Schenk and others (2010))

The systems-based data model based on SYSMOD implements the aforementioned constructs into its core data structure. It extends this core structure with entities supporting the different data formats: numeric values, texts, files, geometries and lifetimes. The core of the resulting data model obtained by Schenk and others (2010) is presented in figure 1. It provides a data model generically able to handle any systemically organised data and therefore potentially any kind of environmental data.

# The SIS architecture

The systems-based information system (SIS) has an architecture known as 3-tier architecture. The first tier is the database, the second tier is a webservice, and the third tier is the client application.

The database is an implementation of the systems-based generic environmental data model introduced in the previous section. It was deployed on a server operating the open-source database management system (DBMS) PostgreSQL (PostgreSQL Global Development Group 2009), with the PostGIS (PostGIS Community 2009) extension to manage the geographic data.

The webservice is a software residing on another server. It acts as an intermediate agent checking requests from the client application and securely accessing the database. On successful identification from the client, it sends back the requested data.

The client is the front-end software containing the graphical user interface (GUI) and the logic to request and process the data from the webservice. Whereas the first and second tiers reside on servers, the client can be installed and used by different users who thereby share the database resources.

The software was programmed in C# and uses the .NET and WPF (Windows Presentation Foundation) frameworks. Its use is therefore restricted to the Windows platform.

# Using the SIS

The client user interface (figure 2) is divided into two main parts. The various tools and menus are in the left panel and the informative, work views are in the central, main window. To access the content of a database, the user enters identification credentials. Once connected, the client downloads the class elements (node, information and interaction classes), which populate the three libraries in the left panel. The class elements in the libraries can then be used as templates. By dragging and dropping them onto the system diagram surface (4), new instances or group instances can be created. For example, the “First reservoir” is a node (rectangle shapes), and is thereby an instance of the “Reservoir” class. “Little town’s reservoirs” is a node group (stack of rectangles) and is thereby a group instance of the “Reservoir” class. Similarly, information classes from the information library (2) can be dropped as information groups (stack of circles) and interaction classes (3) can be dropped as interactions (simple arrows). Furthermore, the interactions library contains five subtypes of “Connections”, which reflect the different structural relations defined in the database. For example, the “Group member” connection was dragged and dropped onto the system diagram surface to specify that “Second reservoir” is a member of the “LittleTown’s reservoir” group.

Instead of creating new instances or group instances, there is the possibility to add the class itself to the diagram surface (5) (the choice between the three forms – class, group, instance - being done by using a key modifier). Classes and instances can be linked by instantiation connections (e.g. “Second reservoir” is an instance of “WS Reservoir” -WS standing for water supply).

The data related to the system diagram elements can be consulted and edited in the Data Catalogue (6). Reflecting the database genericity, the Data Catalogue interface allows handling multiple data formats: numeric values, texts, files (e.g. pictures, videos, pdf files), lifetimes and geometries (in well-known text format). For example, in figure 2 the upper Data Catalogue window shows the data related to the “Second reservoir”. This includes a group of information called “$2^{nd}$ res. water flow” (this information group is linked to the “Second reservoir” node through a property connection in the system diagram). The latter is a water flow indicator whose format is a numeric value. It has several members which are the actual values being displayed in the Data Catalogue. These values can also be used in the system diagram view: by associating numeric values to fluxes, arrows with proportional thicknesses can be drawn (such as the flux between the first and second reservoir).

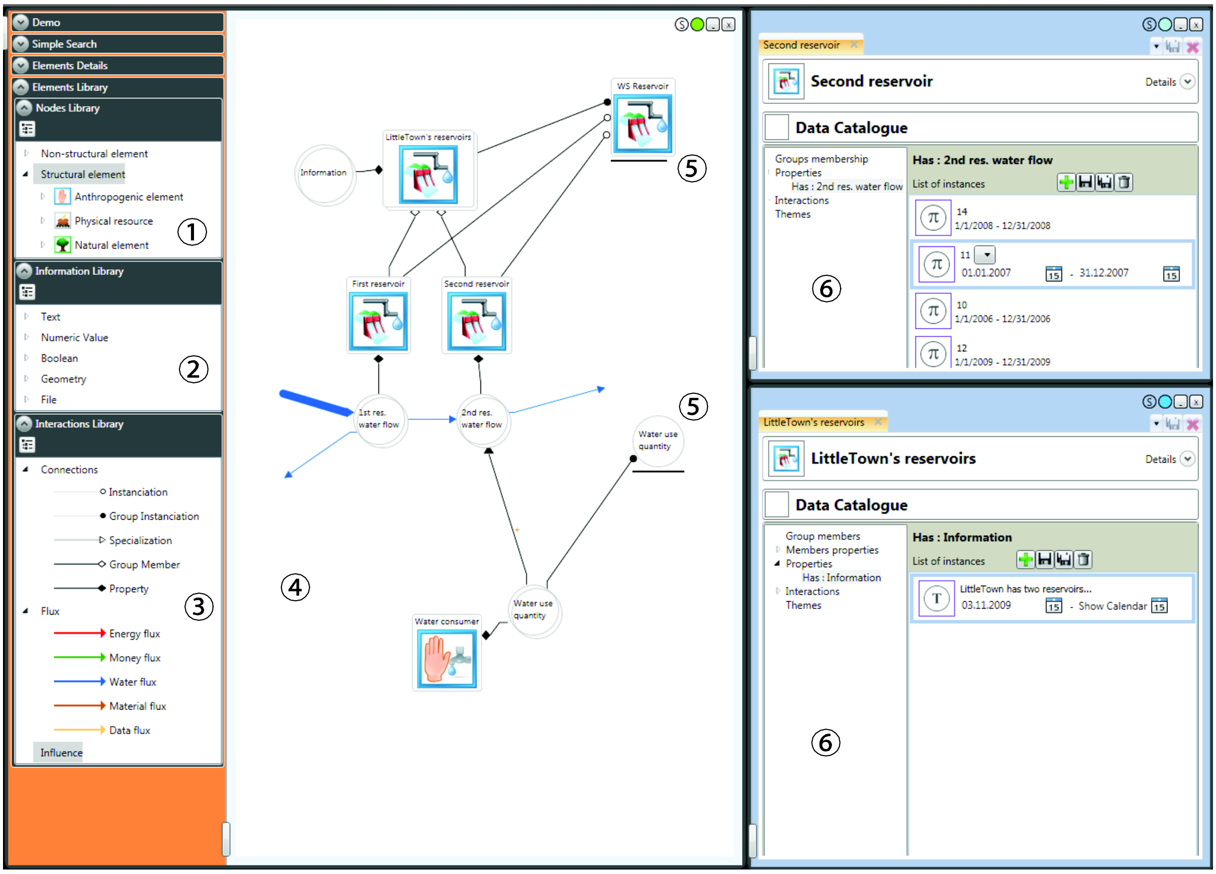



Figure 2: SIS user interface overview

An alternative way of using the system view is the "exploration mode", where elements are dynamically fetched and laid out. This is illustrated by the figure 3, which shows in the upper left view a single starting element, "SWITCH researchers" (from one of the application case studies, used also in figures 4 and 5). By clicking on this element in exploration mode, the attached properties, classes and instances are fetched from the database and displayed in a radial way (lower left view). Then, other elements can be further fetched and displayed (upper and lower right views), allowing exploring the network of interconnected data dynamically. This exploration mode is also useful to manage the data structure. Whereas static, thematic views allow focussing on a given aspect, dynamic views are well-suited to display and edit the whole, various interconnections of elements.

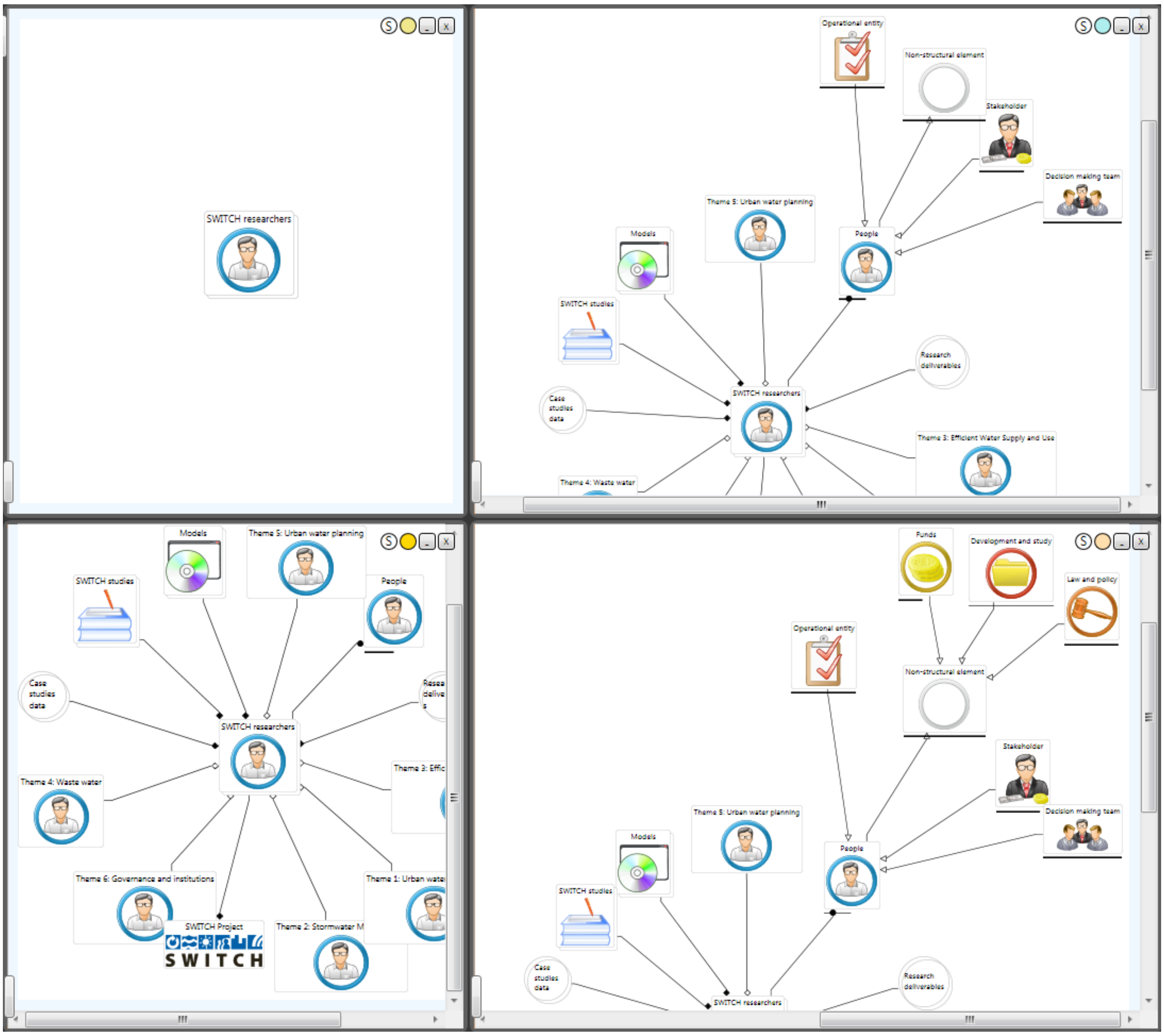



Figure 3: Dynamic navigation in exploration mode

# Application

The SIS was first applied to represent the SWITCH project organisation and handle its data and then to the larger case study of the water management system of Birmingham, UK.

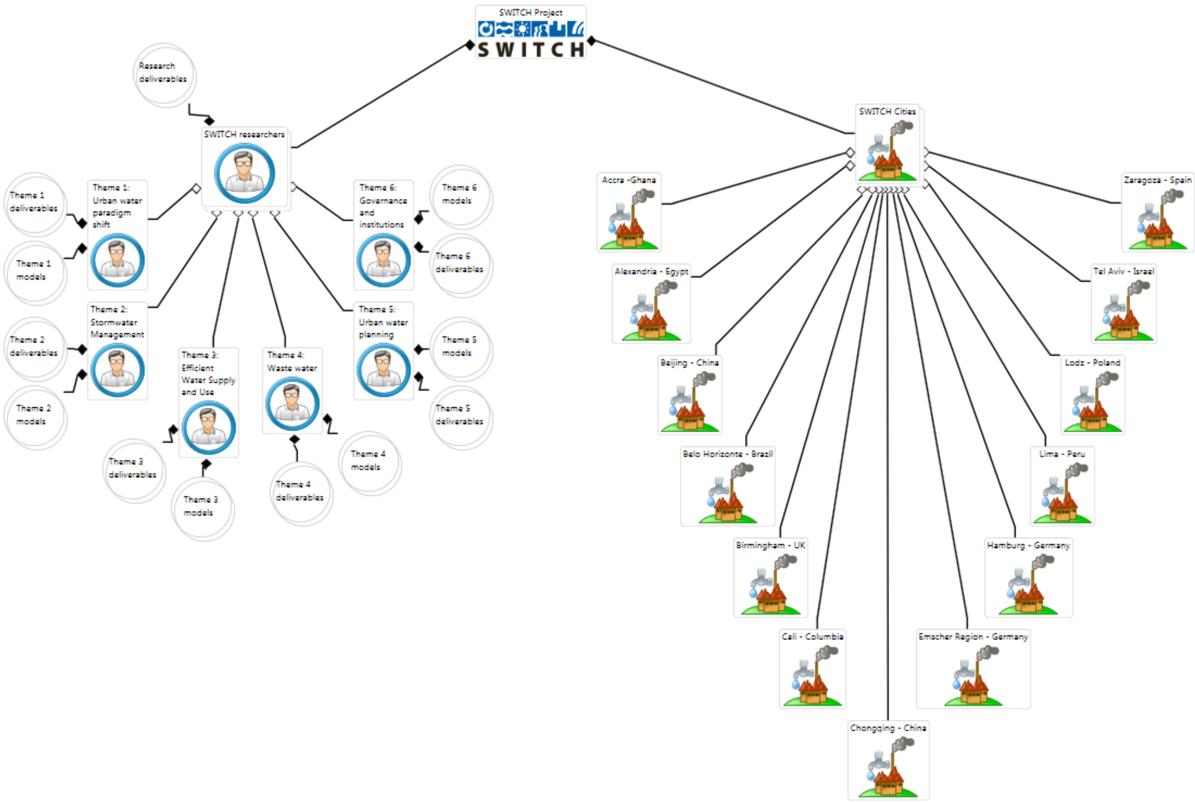



Figure 4: SWITCH project high-level organisational chart

The SWITCH project is about "managing water for the city of the future"; it is a large European project involving many academic teams and thirteen cities all over the world, and funding several research works, including the present study. Its size implies some complexity when it comes to coordinate the different partners. Several views were created to provide a global understanding of the project's structure, workflow, and data management. Figure 4 presents the high-level organisational chart of the project. It shows the thirteen cities and the six research themes. Each team produces studies and models, which can be stored in the respective information groups of each theme. This way, it is easy to manage and consult the outputs of each theme. Furthermore, the figure 4 features the alternative data structures capability: each deliverable produced by a theme is not only pointed to by its information group; it is also a member of the overall "Research deliverables". This way, the same data (the actual pdf files) can be accessed either from the bulk set of documents produced within the project, or from the specific theme outputs. Figure 4 is simple for the sake of readability, but more information could to be displayed, such as contact details for the person in charge and the other members in each city and theme, website addresses, or current

activities. Figure 5 presents another aspect of the SWITCH project: the data flows (lighter arrows) at a strategic level. Cities provide their water-related data to the research teams, which use them to produce studies or as model inputs. Studies, as well as the strategic information gained from model outputs, will enrich back the cities' databases. Then, the decision makers are to manage the databases' content and use it to produce strategies. Figure 5 provides a simple view, common for the different cities. Cities' data flow diagrams respect this general scheme, but each one also has its specificities. For instance, Belo Horizonte, Brazil, is more involved with partners dealing with floods and demand management; in Birmingham, UK, there is one contact person who centralises the data requests. Figure 5 also illustrates how the same system elements can be displayed in different ways in different views: whereas SWITCH researchers and the information group "Research deliverables" are linked in figure 4, the deliverables are linked to "SWITCH studies" in figure 5. In the underlying data structure they are linked to both elements.

The other application case study, to the city of Birmingham, is much larger and more complex. It involved the implementation of a full, detailed systemic model for the water management and led to the creation of hundreds of system elements, declined in numerous views with different topics and levels of details, such as governance, ecosystem management, groundwater issue and Upper River Rea floods. Regarding the implementation of the systemic model for water management, this was carried out on the basis of the water management system model proposed by Schenk and others (2009). This model graphically inventories the systemic elements involved in water management, in a broad sense, including related elements such as energy, ecosystems and funds. These inventoried elements were translated into node classes and associated to an icon to enrich their visual aspect. As node classes, they were then displayed in the nodes library in the user interface (number 1 in figure 2) upon connection to the database. Similarly, the interaction classes (populating the interaction library, number 3 in figure 2) were also derived from the analysis of the water management system model: five main categories of fluxes were identified (energy, money, water, material and data). Moreover, a general "influence" interaction was also found to be used widely. Finally, the information classes (populating the information library, number 2 in figure 2) were created on the basis of classic lists of existing indicator sets for water. Once the class elements were implemented, thematic system views (governance, groundwater, etc.) could finally be created: figure 6 shows, as an example, a high-level view over the water cycle. This figure features how actual flow values can be stored into the database through the Data Catalogue and then applied to the flux arrows to

modulate their width. It thereby provides an immediate hint of the proportions at stake, including the importance of the leaks and of the seepage of wastewater into the pluvial drainage network.

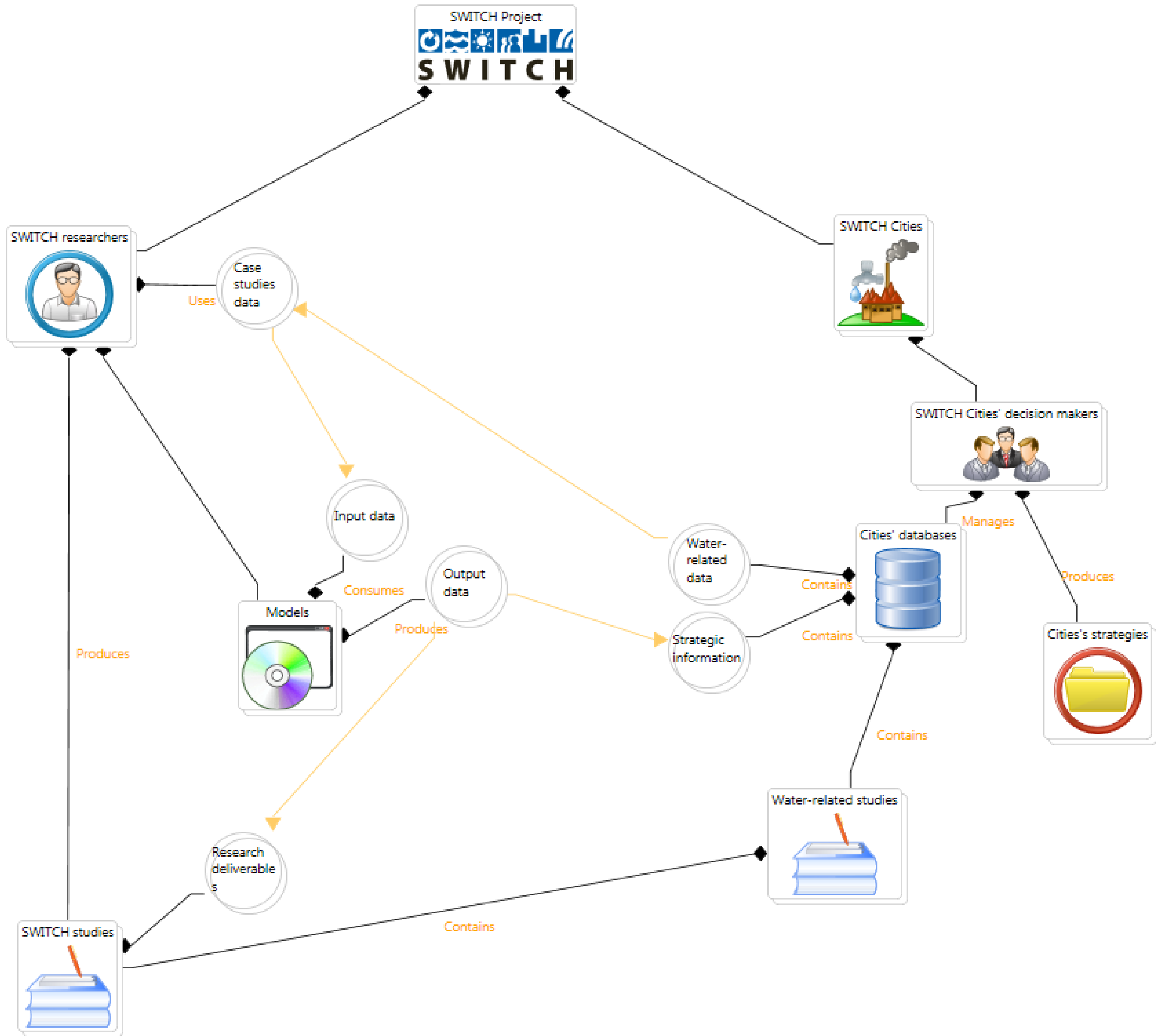



Figure 5: SWITCH project data flows overview

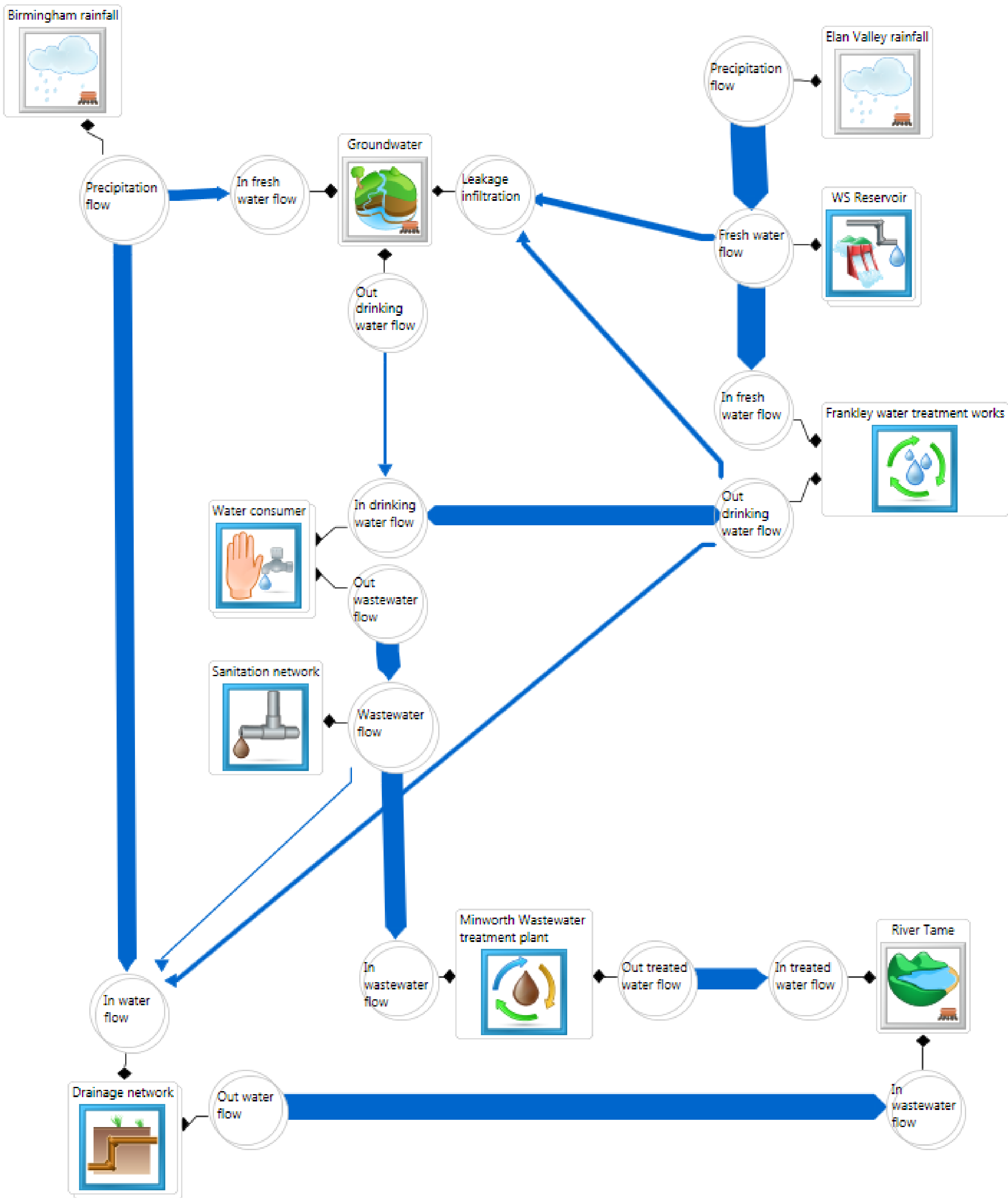



Figure 6: Water cycle overview in Birmingham

To demonstrate the versatility of the SIS, the last example shows a view from a completely different field: an adaptation of the classic example of the roads-and-traffic issue (figure 7). This example illustrates an important aspect of systems, known as feedbacks. Indeed, in the traffic issue view, the

influence relations chain into a closed loop of causes-effects. This means that acting on an element in the loop will result in a retro-effect -a feedback- upon the same element. In the traffic issue case, all the causal relations are positive (+ sign), which means that an increase in the cause results in an increase in the effect. This is known as a reinforcement feedback loop, and it has the property of self-feeding its growth, as an increase anywhere in the loop generates iterative increases after a delay. Namely, in figure 7, an increase in the capacity or density of the road network -for example- results in an increase in the number of people using a car, because traffic was made easier and therefore more attractive. Thereby, the roads' loads increases, which augments the demand for improvements, through more investments into the roads network, to increase its density and capacity. This vicious circle went on and on for decades, and it still goes on, as long as no external factor -such as no more terrain to build new roads- comes into play to interrupt the loop, to limit its growth.

Beside the figures presented, it is important to emphasize that the Data Catalogue always offers the possibility to edit and consult the information linked behind any of the system elements displayed.

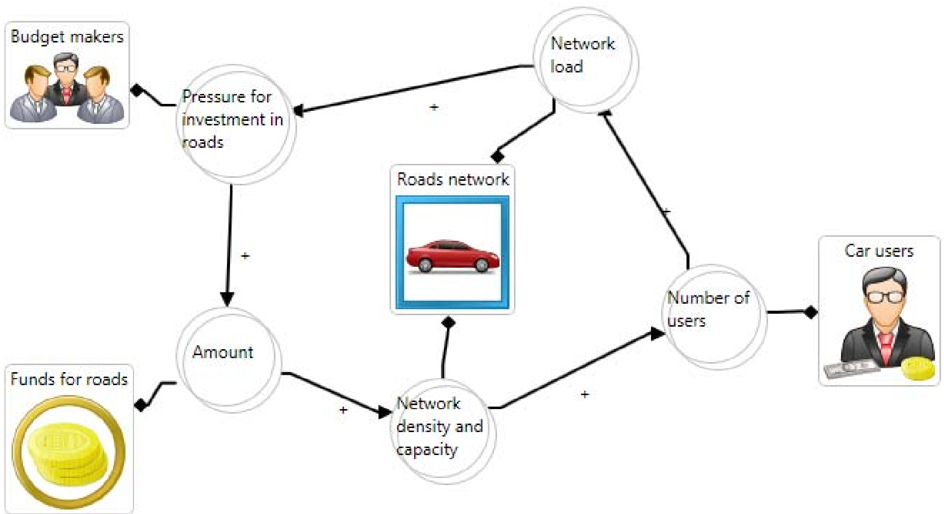



Figure 7: Traffic and roads issue reinforcement feedback loop

# Discussion

Figures 4 to 7, introduced in the previous sections, provide systems-based knowledge of situations. The reason why such knowledge is important is well illustrated by the beer game example.

The MIT beer game (Sterman 1989) involves at least three teams of players: a retailer, a wholesaler and a beer factory manager. The retailer orders beers from the wholesaler, who in turn orders it from the beer factory. At the beginning of the game, a change in the beer consummation happens at the retailer's shop; this is due to an advertisement campaign. To balance its stock depletion, the retailer orders more beer, but because of the delay between order and delivery, its stock quickly nears complete depletion. In the growing fear of empty shelves, the retailer orders more and more. At the wholesaler's store, this increasing demand is multiplied by the number of retailers. Similarly, the wholesaler gradually increases its demand to balance its stock depletion. At the factory, due to the exploding demand, new workers are hired and the process chain is extended, but this takes time and the orders cannot be fulfilled at first. When the factory finally fully reaches its new production capacity, the demand drops to zero. This is because the retailers' and the wholesaler's stocks are now in excess due to their overreaction to the depletion they experimented. In fact, the consummation by the customers did increase at first, but then quickly stabilised. It did not follow an exponential growth. This game has been played thousands of time in different locations with different people, and it always led to similar results: excess stocks and over production capacity at the factory (Senge (2006)). One main reason of this result is the lack of understanding and knowledge of the whole system behaviour by the different teams. While acting separately, trying to make the best for their own situation, they had no clue on what was occurring elsewhere in the system and did not see the global repercussions of their actions.

To avoid problems such as the beer game crisis, systems thinkers advocate a shift of mind, in order to consider situations as systems of elements interconnected by causal relations chained into feedback loops (see for instance Senge (2006), Vester (2007) or Meadows (2008) for recent systems thinking references). They argue that people are generally used to think in linear ways, simple cause and effects, instead of more complex relations with feedbacks. Even the different languages are based on a simple cause and effect structure, with a subject, a verb and an object. Therefore, system thinkers propose causal diagrams as a new language to express more complex, multiple connections with

feedbacks. Such diagrams are then meant to holistically understand a system, and to possibly help pinpointing the levers and bottlenecks.

The SIS fits into this philosophy. The roads and traffic issue (figure 7) is a kind of systems thinking causal diagram, although the formalism is slightly different. However, the SIS goes a step further by the bridge it creates between the systems diagrams and the information systems, since it allows backing graphic system elements with data. Thus, creating views in the SIS such as the SWITCH organisational chart and data flows (figures 4 and 5) or the water cycle view (figure 6) brings several benefits. First, it fosters a holistic apprehension of the situation by proposing a framework stimulating the creator of the view to try thinking to the possible interactions and linked elements; the added visual elements can be documented to provide detailed information (texts, values, files, etc.). Then, the view can be shared (as it is stored in the database), which enables communicating the perception of the situation by the creator of the view. This can be used as a basis for discussion and possibly subsequent modifications. Finally, it provides a reference document and a platform to share information about a given situation.

Beyond the benefits of presenting situations and data in thematic views to communicate them, the SIS is also useful to manage the data. As illustrated by figures 4 to 7, different environmental fields can coexist in the SIS; this is enabled by the generic database. This however requires the development of systemic models for the target fields, as well as their subsequent implementations as class elements. In the case of the water management systemic model, about one hundred classes were implemented. For the traffic issue, the water model was simply extended by adding a new class “roads network” as a subclass of “transport network”. This means that environmental fields’ data can not only coexist, they can be integrated within a single, large systemic model, if the latter is extended to integrate other neighbouring fields (e.g. the water systemic model could be extended to the energy distribution-related data or to the soils-related data). Furthermore, not only complementary models can be integrated, but also alternative models of the same field, which is useful if different needs require different levels of details, or if different stakeholders have conflicting perceptions over a same topic. Here, the key enabling technology is the underlying generic database. But the SIS offers a user-friendly interface for managing the data (especially through the dynamic views, in “exploration mode”).

The SIS is seen as an innovative way of organising data, along its systemic formalism. As such, it may require to get used to it, in order to be able to interpret easily its diagrams. This may be relatively

intuitive for system thinkers, and people used to workflows or organisational charts, but perhaps less for other users, who would probably need more efforts at the beginning. In any case, the SIS is not meant as a replacement of other information systems. In particular, it is not appropriate to visualise maps and in that regard it is rather seen as complementary to GIS. Regarding its performance, it has been used successfully with in-memory datasets of around one hundred system elements (as an order of magnitude), but beyond this size, and depending on the internet connection and the computer, the transaction times with the database start to detract the comfort of use. However, this doesn't affect the size of the datasets in the database, and anyway, adding too many elements into a view finally results in cognitive overloads. In conclusion, the SIS is considered an appropriate solution for institutions or enterprises wishing to integrate, centralise, various datasets in various formats, and for which the systemic structure is relevant in order to create and share views over possibly complex and interconnected situations or problems including: organisation charts and responsibilities, production chains, workflows, catalogues of components and their composing parts, natural resource management, transport networks, etc.

# Conclusion and perspectives

Existing environmental information systems are mostly geographic information systems (GIS) or mere data warehouses. The former imply a geo-centrist data organisation, and the latter are designed for specific uses and lack an advanced, user-friendly interface. In this study, an innovative systems-based information system (SIS) is proposed. It uses a systems-based database which allows generically handling any environmental data modelled systemically, and various formats including numeric values, texts, and multimedia files. The SIS client's first benefit is the ability it provides to consult and organise the datasets in the database in a user-friendly way. The second benefit lies in the possibility of creating thematic views to fathom, document, share and communicate a systemic view of a situation with data behind its elements. Therefore, the SIS is especially relevant for institutions or enterprises dealing with large, various datasets and complex situations which can be expressed as a system view. This could apply for instance to a car manufacturer, whose operation may involve a complex production workflow with the production of parts and their combination at

several different sites (or by sub-contractors), stocks and flows, many people and services with different skills and responsibilities, development strategies with selling figures, research and development projects and results, and a customer service providing feedbacks. With the SIS, the car manufacturer could centralise and manage its various data in one database. It could create views such as organisation charts, process workflows, suppliers' network, causal diagrams of a new marketing strategy, etc.

In the future, the SIS's usability could be improved by further developments, especially regarding the management of geometries and time series. Indeed, the SIS cannot display geometries or maps. To overcome this limitation, it could be extended by a geographic module which would deal with the particular spatial data format. The Data Catalogue is handy to consult a few numeric values, but it is not appropriate to display long time series. For this, a chart module could be added. Besides the possible useful developments of the software itself, the SIS is seen as first step towards a more fundamental way of organising and managing the data, which enables very interesting further research possibilities. The ability to organise the same data in various, possibly alternative ways could be harnessed to export data with various structures, with respect to external needs. For instance, hydrologic data (e.g. water catchments, rainfall series or streams topology) could be organised into several co-existing systems corresponding to the data structures required by different models to run simulations. Thereby, the SIS could support effectively the interoperability of models, and in a more general way, data exchange. Another interesting perspective is the integration of existing systems-based tools or methodologies to perform analyses such as stocks and flows simulations, feedback loop and bottlenecks identifications, or Bayesian networks calculations.

# References

Comunian A, Renard P (2009) Introducing wwhypda: a world-wide collaborative hydrogeological parameters database. Hydrogeology Journal 17 (2): 481-489.

Cooper ER, Siewicki TC, Phillips K (2008) Preliminary risk assessment database and risk ranking of pharmaceuticals in the environment. Science of the Total Environment 398 (1-3): 26-33.

Jorge LAB (2009) Soil erosion fragility assessment using an impact model and geographic information system. Scientia Agricola 66 (5): 658-666.

Kavvadias KC, Maroulis VZ, Krokida MK, Rahman MS (2009) Development of a Versatile Database for Transport Properties of Foods. International Journal of Food Properties 12 (3): 522-535.

Klass JI, Blanford S, Thomas MB (2007) Use of a geographic information system to explore spatial variation in pathogen virulence and the implications for biological control of locusts and grasshoppers. Agricultural and Forest Entomology 9 (3): 201-208.

Li HY, Shi Z, Cheng JL (2007) ARGIS: an Agricultural Resource Geographic Information System for site-specific management of reclaimable saline soils. New Zealand Journal of Agricultural Research 50 (5): 813-821.

Meadows DH (2008) Thinking in Systems. Chelsea Green Publishing, White River Junction, Vermont.

Miege C, Choubert JM, Ribeiro L, Eusebe M, Coquery M (2009) Fate of pharmaceuticals and personal care products in wastewater treatment plants - Conception of a database and first results. Environmental Pollution 157 (5): 1721-1726.

Nishihama S, Hato T, Sugiura D, Yoshizuka K, Sakurai Y (2007) Development of web environmental geographic information system for evaluating river water environment. Bunseki Kagaku 56 (4): 249-253.

Peredo-Parada M, Martinez-Capel F, Garofano-Gomez V, Atenas M, Riestra F (2009) Eco-hydrological database of chilean rivers: a tool for management of aquatic ecosystem. Gayana 73 (1): 119-129.

PostGIS Community (2009) PostGIS. Accessed online 01 / 09 / 2009: http://postgis.refractions.net/.

PostgreSQL Global Development Group (2009) PostgreSQL. Accessed online 01/09/2009: http://www.postgresql.org/.

Pradhan B, Lee S, Mansor S, Buchroithner M, Jamaluddin N, Khujaimah Z (2008) Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model. Journal of Applied Remote Sensing 2.

Roquier B, Schenk C, Soutter M, Mermoud A (2010) SYSMOD: A systems modelling language for environmental information. Submitted to Ecological Modelling.

Sander J, Andersen E, Athanasiadis IN, van Ittersum MK (2009) A database for integrated assessment of European agricultural systems. Environmental Science & Policy 12 (5): 573-587.

Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A (2010) A systems-based, generic environmental data model. Environmental Management Submitted.

Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management. Environmental Management 43 (3): 458-469.

Senge PM (2006) The Fifth Discipline: The Art & Practice of The Learning Organization (revised edition). Broadway Business, 445 pp.

Shih YC, Chou CL, Chiau WY (2009) Geographic information system applied to measuring benthic environmental impact with chemical measures on mariculture at Penghu Islet in Taiwan. Science of the Total Environment 407 (6): 1824-1833.

Soutter M, Alexandrescu M, Schenk C, Drobot R (2009) Adapting a geographical information system-based water resource management to the needs of the Romanian water authorities. Environmental Science and Pollution Research 16 (S1): 33-41.

Sterman JD (1989) Modeling managerial behavior - misperceptions of feedback in a dynamic decision-making experiment. Management Science 35 (3): 321-339.

Vester F (2007) The art of interconnected thinking. MCB Verlag, Munchen, 365 pp.

# CHAPTER 4

## A GENERIC DECISION SUPPORT SYSTEM

# Preface

This chapter describes the generic decision support system (GenDSS), which is the final product of the thesis, leaning on all previous developments. The GenDSS is composed of a database whose data model is defined in Chapter 2, and which itself leans on the SYSMOD systemic formalism (in Annex I). It extends the systems-based information system (Chapter 3) with a fully-fledged user interface and additional modules including the dynamic reporting tool (Annex II) and the data exchange and modelling functionalities (Annex III). It is tested with an implementation of the water management system model, defined in Chapter 1.

This chapter is adapted from a paper in preparation, written by Schenk C, Roquier B, Soutter M, Mermoud A, and which will be entitled: A generic decision support system.

# Abstract

Decision support systems (DSS) are generally designed for specific tasks. Integrated modelling frameworks (IMF) propose frameworks helping creating DSSs, but are not optimal for managing and sharing heterogeneous data, which is rather the focus of information systems (IS). The latter are often mere data warehouses or geographic IS (GIS) with a specific, fixed data structure which doesn't allow easy modifications. Furthermore, GISs are not well suited to display rich non-spatial information. The present study proposes an innovative framework, which bridges ISs and IMFs into a tool able to manage heterogeneous, transdisciplinary data within a versatile central database and display it within multiple complementary views, and where external models can be linked. This tool uses a systems approach which provides a powerful and flexible basis, making it generically applicable to various data and situations. Applied to a given situation, it can provide help to decision making; it is therefore called a generic decision support system (GenDSS). Its application to the case study of the Upper Rea catchment in Birmingham (United Kingdom) demonstrates its potential as an integrated solution for managing heterogeneous data and complex situations, including various scenarios.

**Keywords:** systems approach, data integration, integrated management, information system, decision support system, integrated modelling

# Introduction

A decision support system (DSS) can be defined as an information system topped with tools performing analyses to help making informed decisions. Multitudes of such DSSs have been developed in the past decades and are still being created, generally aiming at a specific situation or sometimes at a specific kind of problem. For instance, in 2009, dozens of DSSs were applied to the field of water management only, such as the works of Gastelum and others (2009) on the management of the Conchos Basin, Hadihardaja (2009) on reservoirs and sediment management, Halide and others (2009) about aquaculture, Kardel and others (2009) in the Biebrza national park, Monte and others (2009) targeting radionuclides and heavy metals, Pedras and others (2009) for irrigation management, Rao and Rajput (2009) for canal management, Sulis and others (2009) on hydrology in the Caia basin.

DSSs specifically designed for a specific situation are obviously very appropriate but they can't be easily adapted to different problems and their reuse potential is therefore very low (Rizzoli and Young 1997). To remediate this issue, various integrated modelling frameworks (IMFs) offer suites of software engineering tools that allow a more rapid development of DSSs, on the basis of reusable components (Rizzoli and others 2008). There are several kinds of such IMFs. Some propose an application programming interface (API) such as OpenMI (Gregersen and others 2007), TIME (Rahman and others 2003), E2 (Argent and others 2009), or ModCom (Hillyer and others 2001). In this case, the DSS developer has to "wrap" the models to be used with programmatic classes provided by the API; then, the different models can be linked, possibly using a graphical user interface (GUI), and run conjointly. Other IMFs propose dedicated programming languages and component libraries, such as MATLAB (The MathWorks Inc. 2009a) or Mathematica (Wolfram Research Inc. 2009) frameworks, which allow the user to develop a DSS (among many other things) using predefined objects such as matrices, algorithms or charts. In yet other cases, the DSS can be composed graphically, using a visual modelling and simulation environment such as STELLA (isee systems Inc. 2009b), Vensim (Ventana Systems Inc. 2009), ithink (isee systems Inc. 2009a) or Simulink (The MathWorks Inc. 2009b).

The aforementioned IMFs are powerful and have been successfully applied in many projects, but they are based on fixed implicit syntaxes. For example, in STELLA, "stocks" and flows" are predefined

as basic elements that can be combined into complex networks, and the users have to accommodate to this syntax. If in two different contexts, the same name is used, but doesn't point to the same concept (e.g. a "resource" representing in one case money and in another case natural goods), it leads to interpretation and integration difficulties. In order to overcome this limitation, some recent works advocate a more semantic approach, based on ontologies. An ontology, in its modern sense applied to computer science, is a formal structure of knowledge. Ontological formalisms (such as OWL, used within the frame of the semantic web) allow defining elements such as classes, instances, and relations (e.g. a student is a subclass of a person). These elements can then be used to provide semantics -the ontology- for a dataset or a model. Consequently, any reader (human or machine) can refer to the ontology to understand what a concept (e.g. the aforementioned "resource" element) really means in its current context. In their recent review, Villa and others (2009) point out a few innovative projects which adopted this new ontology-based approach within their IMF: the SEAMLESS project (Janssen and others 2007, van Ittersum and others 2008) -which proposes an integrated assessment framework for the agriculture, and the Kepler project (Ludascher and others 2004) -which offers a framework to graphically create scientific workflows (workflows involving heterogeneous data and models).

Whereas IMFs' focus is on models and models' data integration, information systems (IS) focus on the integration of information in a larger sense -possibly including comments, documents, pictures, etc.- and its communication. There is currently no application available to act both as an IMF and a generic information system. Most IMFs propose some data visualisation facilities, for instance through simulation results' charts, but even the most advanced ones are not designed for operations such as storing, sharing and displaying pictures or comments by distributed stakeholders through a centralised storage system. Regarding ISs, Schenk and others (2010b) emphasize in a recent study that the multitude of ISs produced continually in the environmental field are often mere data warehouses specific to a given situation, or geographic information systems (GIS) with data structures centred on spatial elements; altogether these were applications with rather specific data structures and model linkages. They argue that a more fundamental and generic way of organising the data is along a systemic language, and therefore proposed a new information system based on such a language. Their systems-based IS (SIS) is in line with the aforementioned ontology-based management of information, as its underlying language SYSMOD (Roquier and others 2010a) can be considered as a kind of ontological language. Thus, the SIS enabled the subsequent development of a module in which systemic definitions of models' inputs could be created in order to link external

models in a generic way (Roquier and others 2010b). The SIS and its subsequent developments therefore led to promising results, but lacked important features, including a proper user interface to manage different views and projects, and the possibility to display geographic elements and charts.

The present study proposes a new software framework, built on the past developments of the SIS and its subsequent augmentations. The new framework provides the missing features of the SIS, namely geographic and charts visualisations, and further integrates all modules within a composite, multi-view user interface. Given its systems-based underlying data structure, it is capable of storing any kind of environmental data (including various formats such as texts, spatial elements, multimedia files, and numeric values) which can then be displayed in various ways: maps, charts, reports and diagrams. Furthermore, the module for linking models makes the framework an IMF. Altogether, this new framework bridges information systems and integrated modelling. As this means that various data and models can be generically handled, it is proposed to define it as a generic decision support system (GenDSS). Thus, the GenDSS is to provide an innovative tool to generically manage and consult data, possibly linking models to run simulations, thereby providing an interesting solution for institutions and enterprises dealing with complex, heterogeneous and transdisciplinary data and situations.

# Integration of available developments

The GenDSS is the result of a long development process, over several years. This process gave rise to several successive products. The theoretical basis lies within the systemic visual language SYSMOD (Roquier and others 2010a). The latter was then applied to the design of a generic data model for environmental data (Schenk and others 2010a). The data model was implemented into a database and a client was coded to retrieve and display the data within a systems-based information system (Schenk and others 2010b). On that basis, a module was developed to propose dynamic reports about system elements (Soutter and others 2010). And another module was designed to exchange data and link models (Roquier and others 2010b). Finally, in the present study, all these

developments were integrated within a fully-fledged user interface and augmented with both a geographic and a chart module.

The visual language SYSMOD (Roquier and others 2010a) defines three types of elements: node, information and interaction. A node is a system element such as a person, a car, a law, a river, etc. It is represented visually by a rounded rectangle. An information element provides values, whose format can be numeric, textual, spatial, a file or a time period. It is represented with a circle shape. An interaction is a relation between two groups, such as a flux or a causal link. It is represented by an arrow. These three types of elements belong to a level, which can be either: class, property, or instance. A class is a template, and an instance of a class is a specific element based on that template. For instance, Birmingham City can be a node instance of the City node class. Classes and instances can have properties which are generally groups of elements. For instance, a car node instance can have a group of four wheels, where each wheel is a node instance of the Wheel class. And each wheel can have a group of information providing it number of punctures. Furthermore, SYSMOD defines five types of structural relations: The relation between an element and one of its properties is called “has property”. The relation between a class and a subclass is a specialisation. The instantiation is a relation between a class and an instance, whereas the property instantiation is the relation between a class and a property. Finally, the group member relation takes place between a group and members of the group.

The database backing the GenDSS implements the SYSMOD constructs into its core data structure (Schenk and others 2010a). Furthermore, it supports various data formats: numeric values, texts, files, geometries and lifetimes. The technology used for the deployment is the open-source database management system (DBMS) PostgreSQL (PostgreSQL Global Development Group 2009), augmented by the PostGIS extension (PostGIS Community 2009) for the management of spatial data.

The first graphical user interface (GUI) linked to the database was called the systems-based information system (SIS) (Schenk and others 2010b) and was designed to visually manage the information of the database and communicate it. With that regard, it offers a diagramming space, where node, information and interaction elements can be structured and connected. Thereby, thematic views can be created, such as organisational charts, workflows, causal loops, or stocks and flows, which are useful to help fathoming the possibly high complexity of the situations they illustrate.

The SIS already offered the possibility to manage the data linked to information groups. It was for instance possible to consult and edit the numeric values of a rainfall time series, or to store the contact details of a stakeholder, within a tool called the Data Catalogue. This functionality was subsequently extended by an active reporting module (Soutter and others 2010) which allowed creating fact sheets for the system elements. Whereas the Data Catalogue presented raw lists of data of a given format, the fact sheets could integrate data in various formats in a single document similar to a web page, and therefore enhanced much the user-friendliness of the communication of the information. These fact sheets can be edited by the user by modifying the content directly, or by selecting and integrating data presented in the Data Catalogue. In the latter case, a change in the data automatically updates the report -therefore making it “active”.

In parallel, a data exchange module was developed to import and export data, and to link models (Roquier and others 2010b). With that regard, the SIS played a key enabling role, as it allows structuring a given piece of data in different ways. For instance, a same document file can be associated to its author in a thematic view, and to the target audience in another view (in the underlying structure, it is linked to both). The data exchange module leaned on that capacity, as it requires the systemic definition of the data structure to be exchanged. For instance, if a model expects inputs such as a river reaches with discharge and temperature values, this structure can be visually modelled as a system into a so-called meta-model. Then, relevant data is selected by the user and marked for exporting. The data exchange model finally produces an XML file, whose structure reflects that of the meta-model. Finally, the model developer has to adapt its product to read the XML input file (if it is not already the case), which is normally a fairly straightforward task, as the data structure is already similar.

# The GenDSS software

The GenDSS builds upon the aforementioned developments. It integrates them into the frame of a multi-view user interface. The GenDSS further provides general tools, such as interfaces for database

connection and project management. It also implements two new modules: a geographic view module and a chart module.

The GenDSS software has a three-tier architecture. The first tier is the database, the second tier is a webservice (providing a security interface for the database), and the third tier is the rich client containing the user interface and the infrastructure code including a data access layer (DAL). The database and the webservice reside on servers. They provide a service which can be accessed by multiple clients, installed on different users' computers, over the internet. To facilitate the communication for casual users, the client can also be opened quickly from a web browser, which doesn't involve a regular installation. The developments were made in C#, using Microsoft's .NET and WPF (Windows Presentation Foundation) frameworks. Therefore, the MS Windows platform is required to run both the webservice and the client. Regarding the latter, which concentrates most of the code, its architecture is modular. This means that the systemic view, the data exchange and the active reporting functionalities are all hosted within loosely coupled modules; removing a module doesn't prevent the other ones to work. This also means that in the future extension modules could be easily added.

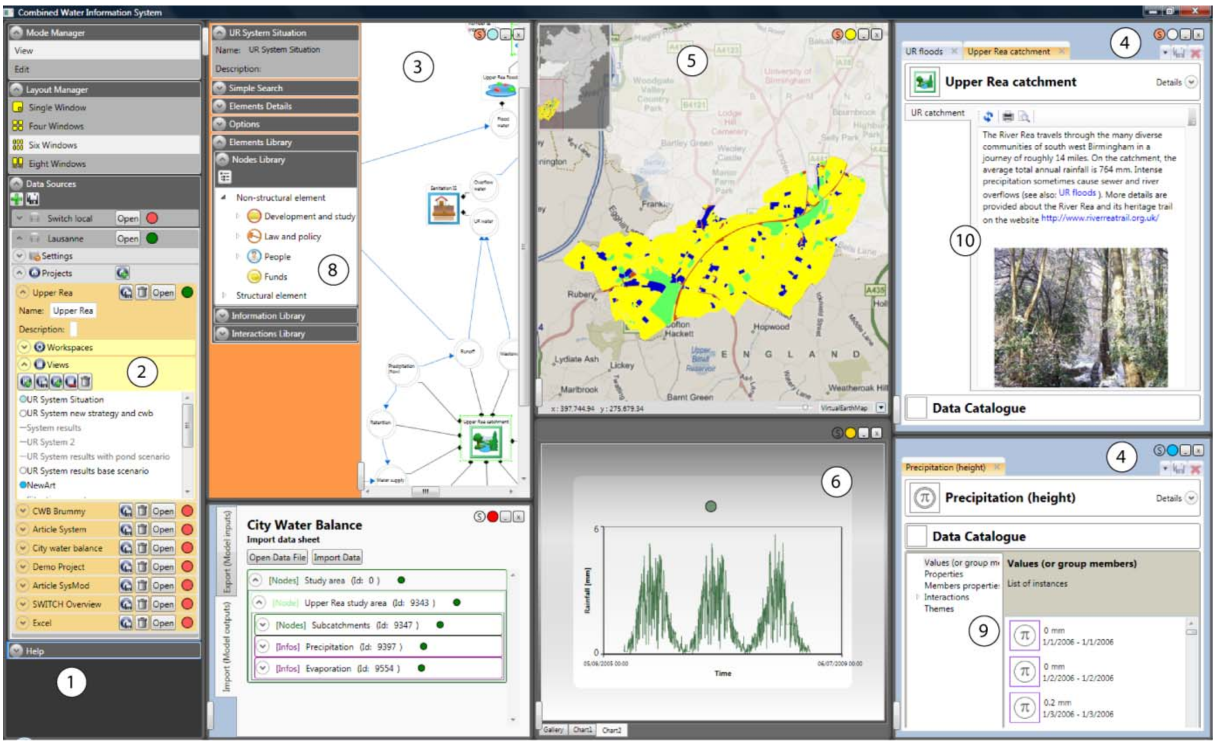



Figure 1: Upper Rea floods issue GenDSS workspace

The GenDSS user interface is presented in figure 1. The general functionalities lie in the left panel (1). They include the possibility to change the layout of the main window, for instance to show a single, large view, or to display several views in parallel, smaller spaces. Another important functionality is the data source toolbox, which allows configuring (URL location, username and password) the access to one or more databases and connecting them. Once connected, the different projects hosted by the database are loaded and listed. They can be opened to gain access to the views and workspace they contain (2). Views are saves of any module's view, such as a diagram in the systemic module or a given map in the geographic viewer. Workspaces are saves of the whole interface's state, providing a "snapshot" of the current GenDSS application window. The right part of the user interface (divided into six smaller windows in figure 1) displays the different open views: a systemic view (3), a data and report view (4), a geographic view (5), a chart view (6) and a data exchange view (7).

The systemic view (figure 1-3) is that of the SIS (Schenk and others 2010b). It visually reflects the systemic organisation of the data, through rounded rectangles (nodes) linked to circles (information groups) by various kinds of arrows (interactions and structural connections). On the left of the diagramming surface, the classes to be used (the underlying systemic model -or ontology) are displayed in three different boxes for the nodes, information, and interactions (8, see also figure 2). By drag-and-dropping class elements, the user can instantiate the classes into instances or groups (groups are symbolised by stacks of shapes), or he can display the class itself (the choice is done by using a key modifier). Arrows' widths can be proportionally modulated by assigning them a numeric value, which allows to represents fluxes as connections with a proportional width.

While the systemic view shows the organisation of the data, the data and report view (figure 1-4) shows the information (texts, values, etc.). Each information group displayed in the systemic view (the circle stacks) can have a number of members which are actual information instances. The latter are not displayed in the systemic view; they appear in the data and report view, as lists (in the "Data Catalogue" -lower part of the view (9)) or as user-friendlier reports (in the "Reports" -upper part of the view (10)). Regarding file information instances (pictures, documents, etc.), they are opened externally by the default associated external application (e.g. a ".docx" file will open up in MS Word).

Another particular case of information instance is the geometry. Its natural way of being displayed is in maps, as points, polylines or polygons, which made it necessary to integrate a geographic viewer (figure 1-5) to deal with this specific case. The geographic viewer provides the basic functionalities one may expect from a geographic information system (GIS): zooming, panning, modifying

geometries' colour and transparency. But where GISs are generally based on data organised in separate tables for different layers, the GenDSS's geographic viewer uses nodes groups. For instance, if a node group gathering the districts of a city is given as the entry point for a layer, the geographic viewer will search and display the district area of each district in the group into a single layer.

Although numeric values can be consulted in the Data Catalogue as a list, there are a number of cases where charts are more appropriate, for instance to compare values (e.g. as histograms) or to see trends (e.g. as line charts). The chart module (figure 1-6) was developed to this end. It is able to display a chart given an input time series. It can also accept as input a group of nodes, to get its node instance members and their associated numeric values (information groups with the numeric format) to automatically propose one or more adequately formatted charts.

Finally, the data exchange view (figure 1-7) can reflect in a tree view the structure of a meta-model (for instance the definition of the data structure a model expects as input). While exporting, the user can populate the different levels of the tree view by drag and dropping data from other views. When dropping an element, it has to match the destination class (the dropped element must have the same class or subclass as the element in the meta-model). In case the dropped element has properties, and these properties match the meta-model definition, sub-levels of the tree view can be automatically populated accordingly, and so on, recursively. Once the data selection is done, it can be exported (data exchange) or the target model (integrated modelling) can be called and run externally and fed with the data from the tree view. Conversely, while importing, the module opens up (data exchange) or retrieves the model outputs (integrated modelling), both cases necessitating a GenDSS-styled XML file. The imported data can be validated by a meta-model (similarly as for the export), which may check the incoming data structure.

The different views not only display complementary information in parallel. An important point of the integration of the different modules consists in the fact that views can be synchronised (this is done by using the "S" tokens at the top right of the views). This means that a selection change in a view provokes the same change in the synchronised views. Thereby, the various facets of a given system element can be visualised in the different views. This is illustrated in figure 1, where a water catchment is selected in the systemic view; the data and report view displays a report about the catchment; the geographic view shows its location, and the chart view shows the associated precipitation values series.

# Application

The River Rea catchment lies in the southwest of Birmingham (United Kingdom). The river is a tributary of the Tame and is historically known for its variability in flows. In the 1700's, the river course was altered to install mills along its course. Subsequently, frequent floods caused significant damages. After a historic flood in 1852, the river was straightened and deepened. However, increased urbanisation had subsequent floods further causing important damages to the houses and equipments in the vicinity. The situation was analysed within the frame of the Upper Rea pilot project, led by the DEFRA (Department for Environment, Food and Rural Affairs) (DEFRA 2008), through an integrated approach. This project highlighted the limitation of the current institutional and regulatory framework, pinpointing data sharing and management problems since the River Rea floods-related data is not centralised and gets lost. It led to the creation of an integrated flooding model (considering both sewer and river overflows), explored scenarios, and provided some general recommendations.

The application of the GenDSS makes use of the outputs of the DEFRA pilot study. It aims at providing a decision support system for the Upper Rea catchment, through the documentation and communication of the situation (information system aspect) topped with the linking to a scoping model (integrated modelling aspect).

To analyse the River Rea situation, an underlying systemic model (an ontology) was required, to define the classes (the templates of the elements) to be used. In their development, Schenk and others (2010b) proposed such a systemic model, based upon a graphical definition elaborated in a previous work (2009). This systemic model was applied to the Upper Rea case study, as it adequately covers the field of water management in a broad sense, including elements such as energy or monetary aspects.

The River Rea situation was then analysed, data was collected and system elements were instantiated using the underlying systemic model classes. The systemic elements were structured into a network, and data was associated to these elements, including geometries, texts, files and values. Selecting elements out of this rich documented network, various views were created. Some representative views were then chosen to be part of the "Upper Rea floods issue" workspace, shown

in figure 1. The various views of this workspace provide complementary information. The geographic view shows the land use of the catchment, divided into small homogenous areas (each area -also called “cluster”- is a system element). The report view provides a fact sheet about the Upper Rea catchment and its flooding issue, including a description, a precipitation value and a picture, as well as a link to another fact sheet and an external link to a website, based upon information stored within the Data Catalogue (both the data in the catalogue and the report are associated to the “Upper Rea catchment” system element). The chart view shows a one-year rainfall series (the data being a series of information, members of the precipitation information group linked to the “Upper Rea catchment” system element). The systemic view is enlarged in figure 2. Its purpose is the documentation and communication of a possible perception of the historic situation regarding floods: heavy rainfalls run off through the catchment (water fluxes). They may cause sewer overflows, or river floods. This historically triggered a “vicious circle” (a reinforcement feedback loop with “influence” arrows, featuring a “+” or “-“ sign to indicate whether the change on both side is proportional or inverse): damages done to estates caused people to claim a better protection. Traditionally, protection was achieved through heavy structural measures, such as deepening and straightening the water course. These equipments provided a somewhat false sense of security, thereby fostering more construction in vulnerable zones. Subsequent stronger floods (with a rare return period, possibly occurring more often in the future due to climate changes) caused then even more damages. This perception of the situation, although questionable, highlights the non-sustainability of the past and current practice. By pinpointing the traditional flood management approach as a potential point to break the vicious circle, it calls for a change in the approach, rather than for pursuing the historical trend consisting in more investments into traditional infrastructure, thereby further feeding the loop.

A new, potentially better integrated approach to the floods issue, breaking the feedback loop, is proposed in a new view, shown in figure 3. The traditional approach is replaced by the integrated approach, as partially experimented by the DEFRA study. This new approach would strive for integrating various data into a central database, data that could be used by models to produce simulations for studies. The studies’ outputs would be used within a participatory framework, favouring the communication of results, and therefore sensitising inhabitants to the floods-related risks (“-“ influence, breaking the positive feedback loop). They would be also used to provide sustainable measures against floods, based on an integrated apprehension of the situation. These measures could possibly take place upstream, intercepting the rainfall or the runoff (e.g. with green

roofs or other retention locations), instead of the traditional measures taking place downstream, and consisting in modifying the river and its embankments to pass always bigger peak flows.

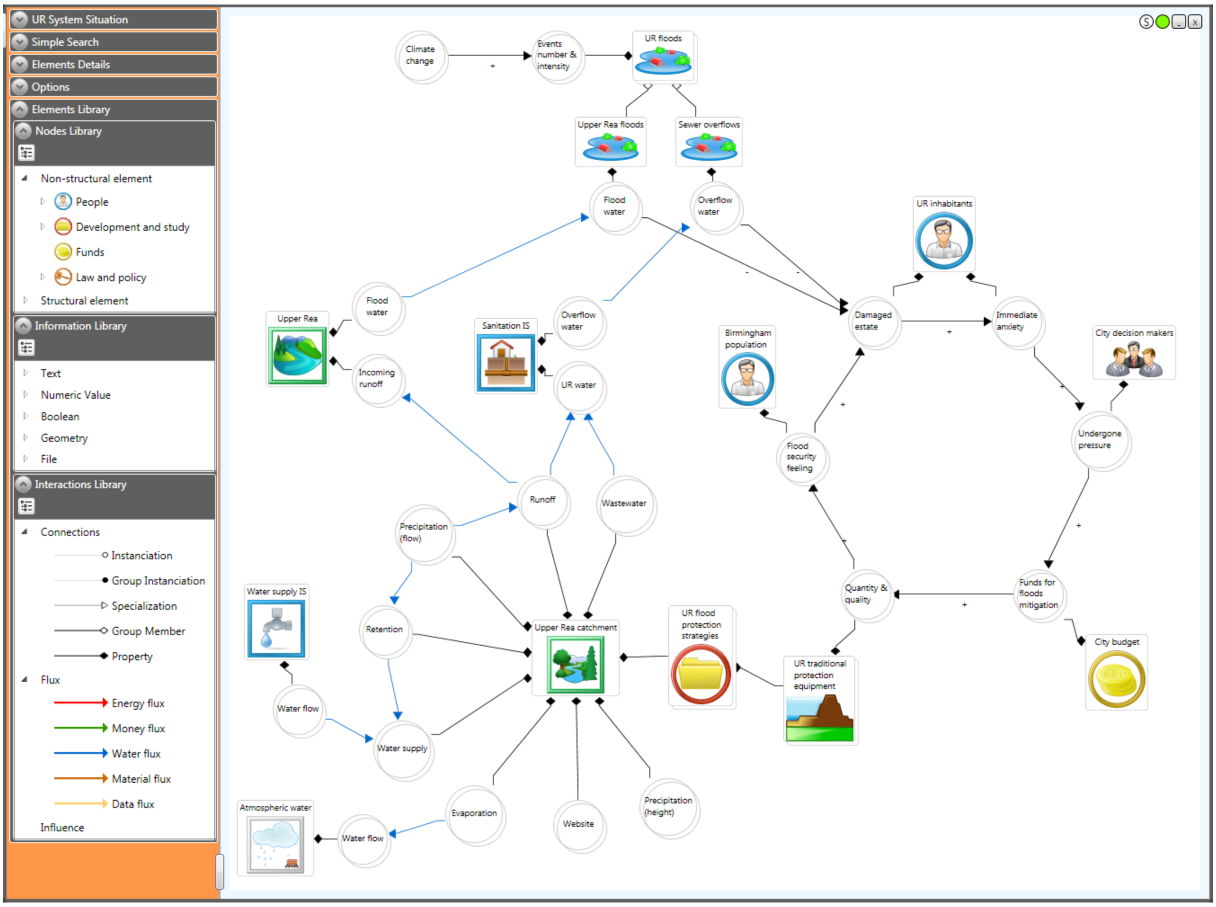



Figure 2: Upper Rea floods issue systemic view

A model is required to predict the impacts of such new measures. The model linked in this specific case was City Water Balance (CWB), developed by the University of Birmingham (Last and Mackay 2010). The linking process was performed and described in details by Roquier and others (2010b), and is briefly explained hereafter. CWB requires the definition of a study area, linked to catchments themselves divided into clusters. These clusters must have a geometry and have their land use defined by a member of the so-called "Unit Block" typology. This organisation is defined in a meta-model view, which is simply a systemic view playing the role of declaring a data structure. The CWB inputs' meta-model view is shown in figure 4 (on the top-left, systemic view). The "Unit Blocks" typology was defined as an extension of the underlying water management systemic model, thereby creating new "Unit Block" node classes. This allowed instantiating this new class into the unit block

element linked to the clusters in figure 4. On the basis of the inputs' meta-model, data can be exported to the model as an xml file reflecting the meta-model structure. Although this can be done manually by drag-and-dropping elements into the export view (on the top-right of figure 4), there is also the possibility to have the selection algorithm automatically populate the export list; but this requires that the data have a compatible structure. To do this, the original data structure, shown in figure 2, was to be augmented. A study area was to be added, and linked to precipitation data. The latter already existed as an information group, property of the "Upper Rea catchment" element; it was simply also linked as a property of the newly added study area node (the result is visible in figure 3). Also, each cluster needed a link to a unit block, requiring the creation of new links.

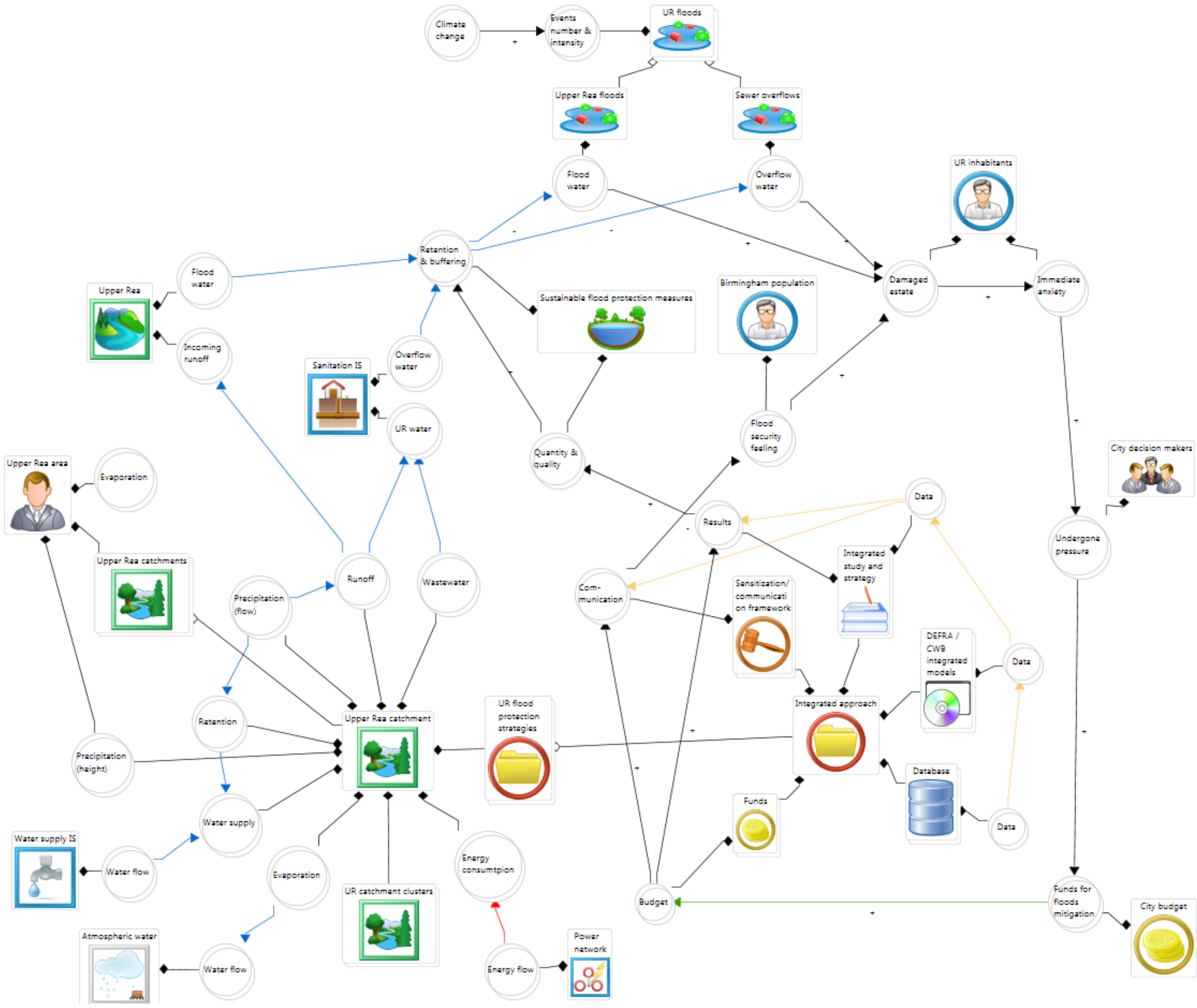



Figure 3: Managing Upper Rea floods with an integrated approach

Once the data structure was made compatible, it allowed quickly exporting the data, by drag-and-dropping the study area element (figure 4, bottom-left systemic view) onto the data modelling and exchanging view (figure 4, top-right view). The algorithm looked for compatible elements in the data structure and automatically populated the tree of data to export, as shown in figure 4.

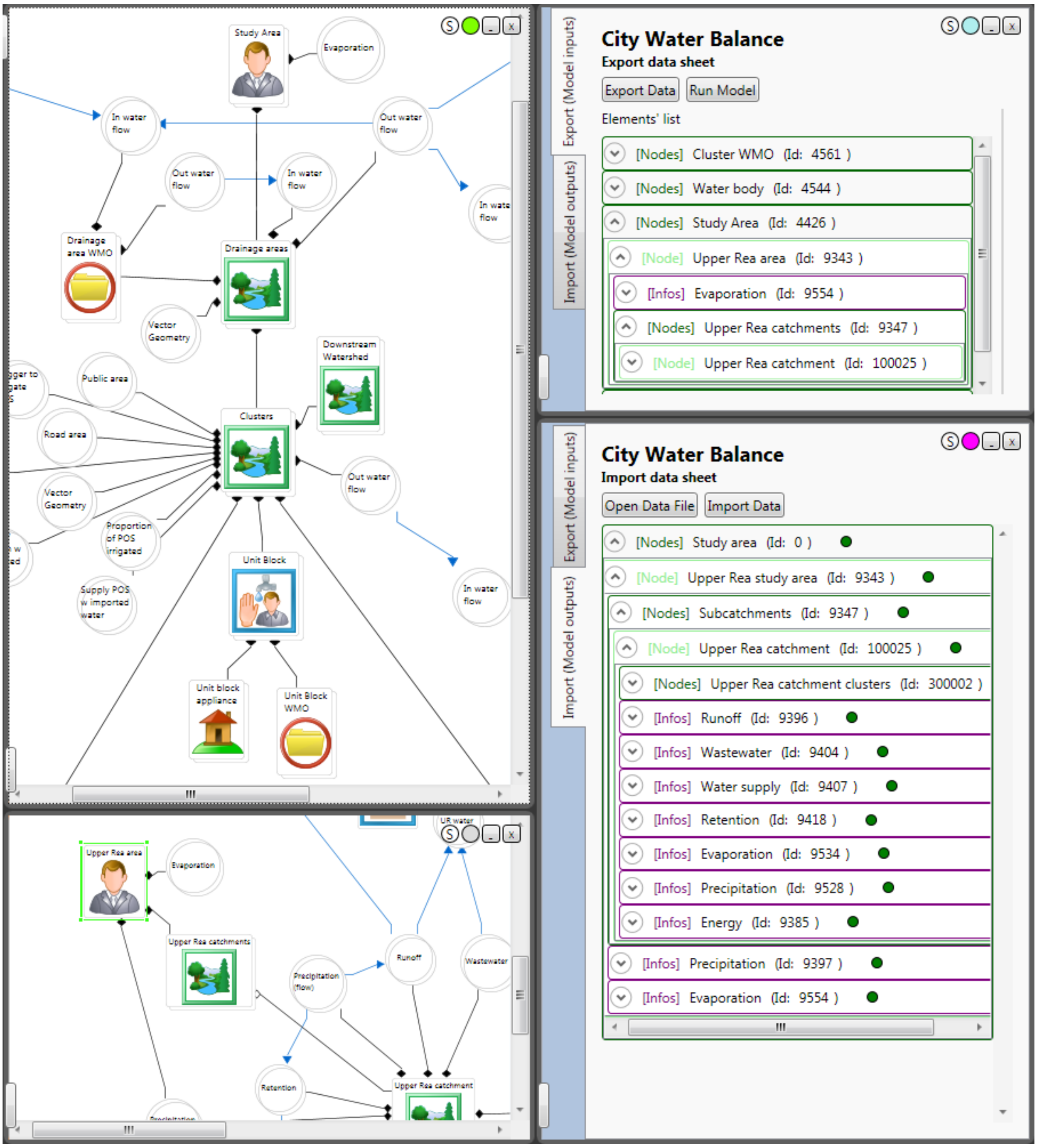



Figure 4: City Water Balance inputs' meta-model, export and import views

Subsequently, the data were passed to the CWB executable, which can be called from the GenDSS. Upon completion, the outputs of CWB were parsed by another piece of software which transforms the output format of CWB into GenDSS-compatible xml. Finally, the results were displayed in the import panel (figure 4, bottom-right). Once reviewed and validated, these results could be imported as new GenDSS elements. For each element, the import process looks for previously existing matching elements to update them with the newly imported values. If no existing element is found, a new one is created.

Using CWB, different scenarios could be tested and compared. This is illustrated by figure 5. In the systemic view, at the bottom, a new retention pond was added. It was allocated a geometry, and linked to two clusters with rather impervious surfaces (the runoff water from MC11674 and MC11735 flows into the new retention pond). The results of this new configuration, as simulated by CWB, were used to create the thematic geographic view, and to modulate the water flows in the systemic view (water flow arrows with a proportional width), as well as the energy consumption (on the right of the view). Finally, the chart view compares results of the pond strategy with the base scenario. The small retention pond was found to have a very limited, but visible effect.

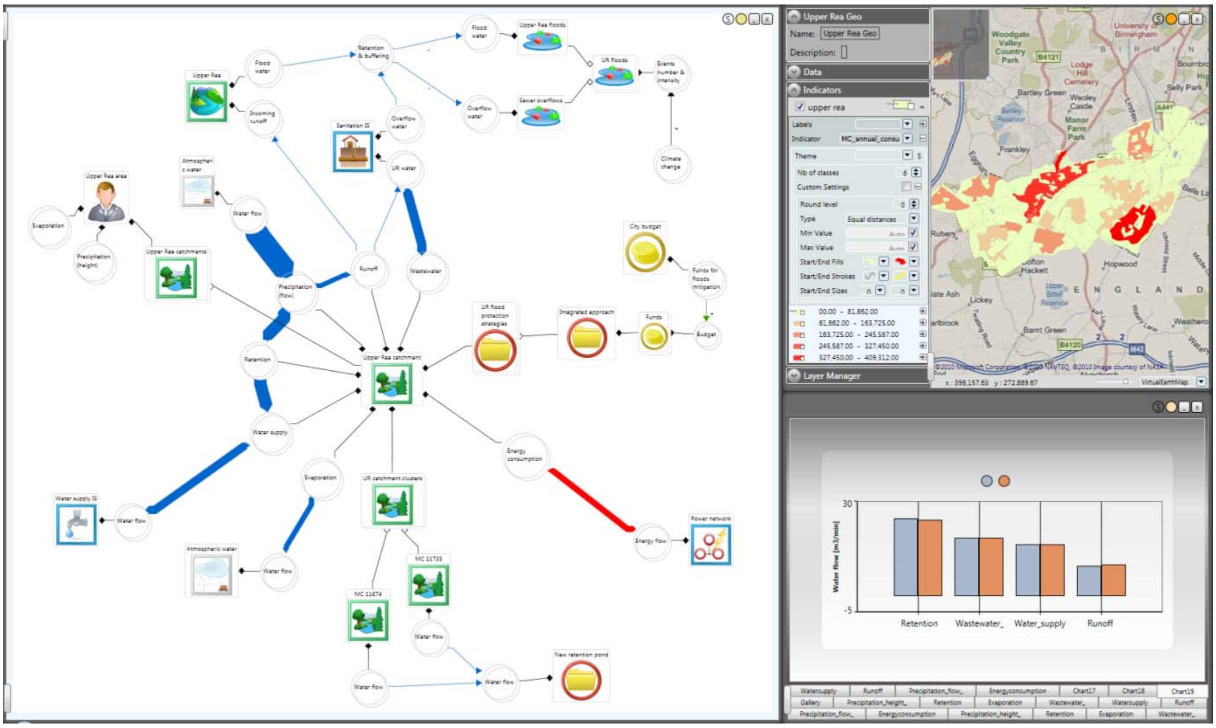



Figure 5: Pond strategy results and comparison with the base scenario

# Discussion

The Upper Rea case study demonstrates the potential of the GenDSS. The different aspects of systemically organised data can be displayed in various, complementary ways. Thereby, a situation - as perceived by the composer of the views- can be described holistically, the system approach offering a powerful and flexible way to integrate various elements and define their interactions. The described situation can be augmented by alternative views reflecting other perceptions, including different scenarios, and these various perceptions can be documented, possibly using linked models. Thus, the GenDSS provides an innovative, centralised solution to integrate and manage heterogeneous, transdisciplinary data and models. It works as a communication platform, as various stakeholders can consult and possibly edit the data from the central database. In summary, it helps document and communicate situations, simulate and inform scenarios, finally enabling making well-informed decisions. In the Upper Rea case, the communication and exchange of data was stated as one of the barrier to improvement; the GenDSS offers a possibility to break that barrier.

There are also limitations to the GenDSS. Currently, as an academic prototype, it suffers from delays in performing operations on large datasets. For example, working with the six hundreds clusters of the Upper Rea catchment results in loading and exporting times of several minutes. However, when dealing with smaller datasets, the reactivity is comfortable. Besides, some improvements are required to make the user interface more user-friendly. For instance, help pop-ups would greatly improve the usability to beginners. Altogether, some complementary work will be invested at term to make the GenDSS into a truly optimised and robust application.

Beyond its operation limitations, the GenDSS present advantages and drawbacks when compared to existing technologies. Among the range of existing information systems, it proposes a truly innovative way of consulting the information, through its various and complementary views. Regarding decision support systems, specific applications targeting a given situation may allow proposing more tightly-fitting end-products. Nevertheless, such products may not allow easily integrating further features (integration of new data and models). Finally, integrated modelling frameworks may perform better in some situations, but may lack the versatility and genericity of the GenDSS. For example, the OpenMI framework (Gregersen and others 2007) allows iteratively cycling through the linked models for progressing timesteps; the Kepler framework (Ludascher and others

2004) can edit files in order to import or export them in a precisely formatted manner. But neither OpenMI nor Kepler is suitable to manage and exchange pictures or geometries via a shared database.

Finally, the GenDSS is a potential innovative solution for institutions and enterprises facing the challenge of integrating heterogeneous data, and wishing to use a generic and versatile product allowing a continuous evolution. Its current limitations are to be overcome at term. Conversely, the GenDSS may not be appropriate for specific situations, with limited and well-defined types of data, or where further developments and integration of data and models is not planned.

# Conclusion and perspectives

The present study proposes an innovative framework, bridging information systems and integrated modelling frameworks into a tool able to manage heterogeneous, transdisciplinary data within a versatile central database to display it within multiple complementary views, and within which external models can be linked. This tool can be applied to various situations to describe them, using systemic diagrams (which provide a holistic, systems thinking vision), geographic maps, reports and charts. Alternative, complementary views (perceptions) of the situations can be created, including various scenarios. Data can be exported and imported using meta-models, which also allows linking external models. Thereby, various scenarios can be simulated and compared. Altogether, it provides an integrated solution helping performing well-informed decision making.

Therefore, the GenDSS might be of particular interest for large institutions and enterprises, dealing with multiple, complex situations and trandisciplinary data, where continuous developments and evolution are needed. This includes for instance environmental management boards, such as in the Upper Rea case study, or multinational corporations, such as a computer manufacturer. In the latter case, the production workflow could involve different teams, possibly in different countries, with different legal frameworks. The parts could be manufactured and combined at different locations, possibly by subcontractors. Beyond the production workflow, many other teams with different skills could work on research, developments, marketing strategies or at the customers services. The GenDSS could be used by the computer manufacturer to centralise and share its data, using

organisational charts, workflow diagrams, geographic maps of its sites' and customers' locations, charts of selling figures, diagrams of new strategies with possible feedback loops, reports about subcontractors services, and so on. External models to simulate consumption trends, or price changes, within different scenarios, could be linked.

To provide an industrial -robust, reliable and secure- product, the current prototype shall be improved. Beyond the technical aspects, further research perspectives are opened up by the GenDSS. Management and planning aspects, as well as scenarios management features could be added to provide users with the possibility to handle projects, with timeline, tasks and responsibilities, and to compare and follow evolution over time. The systemic structures could evolve into actual "internal" models, with a mathematical behaviour. This could be done through the allocation of behaviours to system elements. Then, an engine able to parse the systemic structure could check its validity, and then use the underlying values to run simulations, without the need to link external models.

# References


Argent RM, Perraud JM, Rahman JM, Grayson RB, Podger GM (2009) A new approach to water quality modelling and environmental decision support systems. Environmental Modelling & Software 24 (7): 809-818.

DEFRA (2008) Upper Rea pilot summary. Accessed online 1/4/2010: http://www.defra.gov.uk/environment/flooding/manage/surfacewater/urpilotrea.htm.

Gastelum JR, Valdes JB, Stewart S (2009) A Decision Support System to Improve Water Resources Management in the Conchos Basin. Water Resources Management 23 (8): 1519-1548.

Gregersen JB, Gijsbers PJA, Westen SJP (2007) OpenMI: Open modelling interface. Journal of Hydroinformatics 9 (3): 175-191.

Hadihardaja IK (2009) Decision support system for optimal reservoir operation modeling within sediment deposition control. Water Science and Technology 59 (3): 479-489.

Halide H, Stigebrandt A, Rehbein M, McKinnona AD (2009) Developing a decision support system for sustainable cage aquaculture. Environmental Modelling & Software 24 (6): 694-702.

Hillyer C, Bolte J, van Evert F, Lamaker A (2001) The ModCom modular simulation system. In: (ed) Proceeding of 2nd International Symposium on Modeling Cropping Systems, Florence, Italy, pp 333-343.

isee systems Inc. (2009a) iThink - Systems thinking for business. Accessed online 11/17/2009: http://www.iseesystems.com/Softwares/Business/ithinkSoftware.aspx.

isee systems Inc. (2009b) STELLA - Systems thinking for education and research. Accessed online 11/17/2009: http://www.iseesystems.com/softwares/Education/StellaSoftware.aspx.

Janssen S, Ewert F, Li H, Athanasiadis IN, Wien JJF, Therond O, Knapen MJR, Bezlepkina I, Alkan-Olsson J, Rizzoli AE, Belhouchette H, Svensson M, van Ittersum MK (2007) Defining assessment projects and scenarios for policy support: Use of ontology in Integrated Assessment and Modelling. In: (ed) Proceeding of 2nd International Conference on Asian Simulation and Modelling, Chiang Mai, THAILAND, pp 1491-1500.

Kardel I, Mirosla-Swiatek D, Chormanski J, Okruszko T, Wassen M (2009) Water management decision support system for Biebrza national park. Environment Protection Engineering 35 (2): 173-180.

Last E, Mackay R (2010) City Water Balance - A new tool for scoping integrated urban water management options. In *1st IAHR European Division Congress, 3rd Int. Junior Researcher and Engineer Workshop on Hydraulic Structures* Edinburgh.

Ludascher B, Altintas I, Berkley C, Higgins D, Jaeger E, Jones M, Lee EA, Tao J, Zhao Y (2004) Scientific workflow management and the Kepler system. In: (ed) Proceeding of GGF Workshop on Workflow in Grid Systems, Berlin, GERMANY, pp 1039-1065.

Monte L, Brittain JE, Gallego E, Hakanson L, Hofman D, Jimenez A (2009) MOIRA-PLUS: A decision support system for the management of complex fresh water ecosystems contaminated by radionuclides and heavy metals. Computers & Geosciences 35 (5): 880-896.

Pedras CMG, Pereira LS, Goncalves JM (2009) MIRRIG: A decision support system for design and evaluation of microirrigation systems. Agricultural Water Management 96 (4): 691-701.

PostGIS Community (2009) PostGIS. Accessed online 01 / 09 / 2009: http://postgis.refractions.net/.

PostgreSQL Global Development Group (2009) PostgreSQL. Accessed online 01/09/2009: http://www.postgresql.org/.

Rahman JM, Seaton SP, Perraud JM, Hotham H, Verrelli DI, Coleman JR, Mssanzi (2003) It's TIME for a new environmental modelling framework. In: (ed) Proceeding of International Congress on Modelling and Simulation, Townsville, AUSTRALIA, pp 1727-1732.

Rao BK, Rajput TBS (2009) Decision support system for efficient water management in canal command areas. Current Science 97 (1): 90-98.

Rizzoli AE, Leavesley G, Ascough JC, Argent RM, Athanasiadis IN, Brilhante V, Claeys FHA, David O, Donatelli M, Gijsbers P, Havlik D, Kassahun A, Krause P, Quinn NWT, Scholten H, Sojda RS, Villa F (2008) Integrated modelling frameworks for environmental assessment and decision support. In: Jakeman AJ, Voinov AA, Rizzoli AE and Chen SH (ed) Environmental Modelling, Software and Decision Support, pp 101-118.

Rizzoli AE, Young WJ (1997) Delivering environmental decision support systems: software tools and techniques. Environmental Modelling & Software 12 (2-3): 237-249.

Roquier B, Schenk C, Soutter M, Mermoud A (2010a) SYSMOD: A systems modelling language for environmental information. Submitted to Ecological Modelling.

Roquier B, Schenk C, Soutter M, Mermoud A (2010b) A systems-based integrated modelling framework. In preparation.

Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A (2010a) A systems-based, generic environmental data model. Environmental Management Submitted.

Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management. Environmental Management 43 (3): 458-469.

Schenk C, Roquier B, Soutter M, Mermoud A (2010b) A systems-based information system. Submitted to Information Systems.

Soutter M, Roquier B, Schenk C, Mermoud A (2010) ART: An active systems-based navigation and reporting tool. Submitted to Information Systems Frontiers.

Sulis M, Marrocu M, Paniconi C (2009) Conjunctive Use of a Hydrological Model and a Multicriteria Decision Support System for a Case Study on the Caia Catchment, Portugal. Journal of Hydrologic Engineering 14 (2): 141-152.

The MathWorks Inc. (2009a) MATLAB - The Language Of Technical Computing. Accessed online 11/17/2009: http://www.mathworks.com/products/matlab/.

The MathWorks Inc. (2009b) Simulink - Simulation and Model-Based Design. Accessed online 11/17/2009: http://www.mathworks.com/products/simulink/.

van Ittersum MK, Ewert F, Heckelei T, Wery J, Olsson JA, Andersen E, Bezlepkina I, Brouwer F, Donatelli M, Flichman G, Olsson L, Rizzoli AE, van der Wal T, Wien JE, Wolf J (2008) Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96 (1-3): 150-165.

Ventana Systems Inc. (2009) Vensim. Accessed online 17/11/2009: http://www.vensim.com/.

Villa F, Athanasiadis IN, Rizzoli AE (2009) Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. Environmental Modelling & Software 24 (5): 577-587.

Wolfram Research Inc. (2009) Mathematica. Accessed online 11/17/2009: http://www.wolfram.com/products/mathematica/.

# Conclusion & Perspectives

The developments completed within the frame of the present thesis strived towards the final objective of proposing a solution enabling a complete integration of various kinds of data and models and their visualisation. These developments were applied to the water management issue as a main case study.

The first development step taken was the development of a systemic description of water management. Albeit the various aspects of water management are abundantly documented and well-known, an integrated description including interconnections was lacking. Therefore, a graphic representation of interconnected system elements annotated with definitions was implemented. This innovative result constitutes an ontology providing a robust reference and a fairly exhaustive reminder of the complexity of the water management field, including transdisciplinary aspects. It has been used in the next development steps as a basis for the application case studies. This work addressed the first research question of the thesis: how to systemically define water management?

In the second step, a data model was created, based on the SYSMOD language. The latter was developed to address the thesis research problem of "what graphical systemic formalism should be utilized?" SYSMOD proposes a new graphical language uniting the ontological and systemic approaches, with a strong emphasis on interactions. The developed data model reflects the SYSMOD constructs, adding support to various kinds of data types, including geometries, numeric values, texts and files. Whereas existing data models are generally specific to a field or designed for a target application, this new data model is generic and transdisciplinary. It can handle any kind of data organised along a systemic structure, which potentially includes any sort of environmental data, thanks to its underlying structure consisting of system elements and interactions. It addresses the thesis research issue of "how to integrate and store various, heterogeneous data?" The data model was a key development, enabling the subsequent work focussing on "how to visualise and manage the systemic, integrated data?"

To visualise and manage the systems-based, generic data stored within the database (created on the basis of the generic data model) a new kind of visualisation was developed, proposing diagrams of interconnected nodes. Whereas existing information systems are often data warehouses without a user-friendly graphical interface, or geographic information systems focussing on spatial data, this new diagram visualisation proposes rich views including spatial and non-spatial system elements (such as lakes, power plants, stakeholders or laws) interconnected by interactions (such as fluxes or influences). Bound to the database, the tool was called an "information system on the system" (ISS).

It provides a way of graphically managing and integrating the systemically organised, generic, transdisciplinary data. Furthermore, the systems approach fosters the apprehension and communication of complex situations in a holistic way (involving many different elements, influences and interactions from various fields), through their representation into thematic views.

In complement to the diagram visualisation, a reporting tool was developed to deal with the actual data linked to system elements: numeric values, comments, files, etc. With this tool, the data can be displayed and edited, and reports can be created, that present synthetic fact sheets in a web-browser fashion. Another module was created on top of the ISS to address the question of "how to leverage the systemic formalism to interoperate with the models?" This module uses systemic diagrams in a role of meta-models defining the inputs' and outputs' data structures of models to be linked. This allows exporting data within xml files with a structure close to that required by the target model; the integration of the latter consisted then in reading and writing the new xml format.

Finally, a fully-fledged general user interface was created to integrate all the previous developments into a single environment. To visualise the different facets of system elements, a geographic and a chart view were added, in complement to the reporting view. This final tool was called the generic decision support system (GenDSS), as its application to a given situation can give rise to a series of views and linkages to models, thereby providing specific DSS functionalities for the situation.

The successive developments were tested in the frame of the case study of Birmingham (United Kingdom). In particular, the final GenDSS was applied quite extensively to the floods issue situation in the Upper Rea catchment in that city. This application highlighted the limitations of the GenDSS. Currently, it is limited to managing a few hundred elements; beyond, the loading times and the user interface can become long and unresponsive. Furthermore, the development of such a complex tool requires a huge amount of time and many improvements are still necessary in the user interface to make it robust and intuitive. Finally, only a few stakeholders really had a chance in the limited time of this thesis to try the Birmingham application, and to provide some feedback. It appeared nonetheless that they needed time and explanations to get used to the rich systemic formalism and views, which suggests some possible acceptance difficulties. On the other hand, the application demonstrates the strengths of the GenDSS which provides an innovative integrated solution to centralise transdisciplinary data in various formats (e.g. numeric values, texts, geometries, files), to holistically describe situations using complementary visualisations, and to test scenarios using externally-linked models.

Although several neighbouring tools exist, none achieves this level of integration. They may perform better in specific situations being within their target application area, but difficulties could arise later, when further transdisciplinary developments or integration with other tools would be required. Thus, the GenDSS may prove an adequately versatile solution for large institutions or enterprises dealing with complex situations, heterogeneous and transdisciplinary data, and where requirements are continuously evolving, such as: a car manufacturing corporation (which would involve several different production and research sites in different countries with different regulations), a hospital, or a strategic environmental planning network (where it could help breaking the transdisciplinary, sectorial barriers); in the case of water management, it could prove very useful to achieve a truly integrated water management approach.

In the future, to fulfil the need for a robust, reliable, well-performing and secure solution for large enterprises or institutions with large datasets, the GenDSS will require technical improvements and optimisations, to overcome its current limitations. Beyond these technical points, the GenDSS requires further testing with various stakeholders to better assess its acceptance, especially regarding the systemic aspects, to possibly modify and simplify it, and to eventually validate it. Regarding further research, it opens-up interesting perspectives. Advanced checking and validation functionalities would prove useful to test the systemic networks created by the users, highlighting inconsistencies. The planning and management aspects could be deepened, in order to provide more advanced tools and visualisations, to deal with different scenarios and various indicators sets, comparing them using multi-criteria methods and showing the results as Pareto fronts. Also, an even further step in integration could be taken through the “internalisation” of models. Instead -or in complement to- using externally-linked models, mathematical behaviours could be allocated to the constituting elements of systemic diagrams, turning the latter into “internal” models presenting the advantage to evolve dynamically, conjointly with their representation.

# ANNEX I

## SYSMOD: A SYSTEMS MODELLING LANGUAGE FOR ENVIRONMENTAL INFORMATION

# Preface

This annex presents SYSMOD, a new systemic language which is used throughout the thesis to structure the data (as explained in Chapter 2) and to express situations (as detailed in Chapter 3).

This annex is adapted from the article: Roquier B, Schenk C, Soutter M, Mermoud A (2010) SYSMOD: A systems modelling language for environmental information, submitted to Ecological Modelling.

# Abstract

SysMod is a new graphical, generic modelling language created to support the development of tools in the field of integrated natural resources management (INRM). To cope with the holistic, interdisciplinary framework imposed by INRM, the language is designed to allow the creation of ontologies that facilitate the integration and the reusability of environment related information.

Based on systems theories, SysMod is built on a list of constructs which are specially defined for modelling complex hierarchical systems. In this sense, the holon terminology is adopted to represent systems hierarchically, the use of the name holon meaning that systems should be considered simultaneously as wholes and parts.

Besides introducing the SysMod constructs, the paper identifies the language requirements towards the development of tools for INRM and assesses the effectiveness of SysMod when compared to other ontology and systems modelling languages. SysMod modelling capabilities are illustrated with a list of short examples.

**Keywords:** environmental data, ontology, modelling language, systems thinking, holon

# Introduction

The things that compose the world can always be interpreted as systems, i.e. some groups of interacting and interdependent elements forming complex wholes. While working in multidisciplinary context, systems approaches are of great interest to deal with complexity and to bring the various fields of science together (Midgley 1992, Mulej and others 2004). This is particularly true in the frame of Integrated Natural Resources Management (INRM) and its related fields, where modelling of the ecological system as well as the social and economic systems are key issues towards sustainable planning (Bellamy and others 2001, Rammel and others 2007). In that sense, INRM requires tools that cope with interdisciplinarity, multiple scales (e.g. spatial, temporal or organizational) and knowledge from various sources.

This paper aims at finding an appropriate language to support the development of modelling tools for INRM. Because of the diversity and the complexity of the subject, such tools may have to deal with many types of models, be they ecological, social or economic, and should possibly integrate them. Besides running simulations, INRM tools may reach their full potential by allowing users to create, consult and share any kind of environmental information, whether in the form of numeric values, texts, spatial geometries or images (Schenk and others 2010b). To handle the aspects mentioned above, the language should match three essential requirements. The first is genericity, as the language should be transdisciplinary, in order to model any kind of system without relying on a specific field of science. The second feature is semantic integration (or semantic interoperability), which relates to the use of ontology. This latter allows sharing unambiguous meaning about the concept used in models. This is particularly important in the context of INRM, where models may be the results of participative processes that involve many stakeholders with a variety of knowledge background. As highlighted by Villa (2009), the integration of ontologies in the field of systems modelling will be an important step towards the reusability and the potential coupling of systems models. The third feature is the ability of the language to model open systems. In contrast to close systems, open systems interact with their environment through the transfer of energy, material or information, making them more difficult to model. As a result, any open system may be considered as part of potentially larger systems.

The main problem in finding a suitable modelling language doesn't lie in the lack of available ontology languages compatible with systems modelling, but rather in the ability of such languages to adequately detail systems interactions. Although this is not their primary purpose, ontology languages may be used for systems modelling, as demonstrated by Tudorache (2006) who explores the application of ontologies for systems engineering. Conversely, systems languages may be used (or extended, in the case of UML) to create ontologies (Kogut and others 2002, OMG 2009). However, both ontology and systems modelling languages fail to successfully model the interactions within environmental systems. On the one hand, regular ontology languages don't have a basic construct for interactions. Although, these latter may be defined through properties or meta-classes, ontology languages are not explicitly conceived for representing flows or influences between elements. On the other hand, in the context of systems languages, an interaction typically provides a single piece of information, whether a specific influence or a flow of a particular kind between two elements, e.g. a flow of data, energy, water or any material. In contrast, we argue that influences and flows should be perceived as possible aggregations of multiple systems elements and/or information. That is the case, for instance, of a water flow which may have a pollutant concentration or an influence that summarizes a complex phenomenon in a single relationship.

This paper introduces a new modelling language that has the abilities to create ontologies and to model the complex features of systems interactions. The paper first details the main requirements of the language for modelling complex systems in the framework of INRM, along with an analysis of existing languages. The new language, called "SysMod", is based on a list of constructs whose organization is transcribed into a meta-model. The use of SysMod is illustrated with the help of short examples about ontologies, systems components and systems dynamics, issues related to systems definition and finally some considerations about time, spatiality and multiple potential scenarios.

# Language requirements

The variety of existing ontology and modelling languages shows there is no single way of modelling systems. According to the arguments brought above, the next chapters will analyse and depict the language features that make a positive contribution to INRM.

**Genericity, holism and holarchy**

It is generally admitted that INRM requires a holistic framework; the general principle of holism being rightly summarized by Aristotle in the: "The whole is more than the sum of its parts" (Metaphysics 1045a10). One example advocating this approach resides in the first principle of the "Dublin Statement on Water and Sustainable Development" which concerns the integrated management of water resources (ICWE 1992): *"Since water sustains life, effective management of water resources demands a* ***holistic*** *approach, linking social and economic development with protection of natural ecosystems. Effective management links land and water uses across the whole of a catchment area or groundwater aquifer."*

By adopting a holistic framework, the modellers, as well as the decision makers or the planners, should therefore accept that the behaviour of a system is not totally foreseeable, no matter how much data are available.

Close to holism, this paper embraces a related perspective in adopting the concept of "holon", a holon being something that is simultaneously a whole and a part (Koestler 1967, 1969). The word "holon" is a composition of the Greek word *holos* meaning "whole" and the suffix "on" implying "part" as in "*proton*". The concept of holon allows the modellers to describe any system as a hierarchy of wholes or, in adopting Koestler's terminology, as a "holarchy", i.e. a hierarchy of self-regulating holons. This hierarchical approach may be of great interest when trying to tackle the philosophical debate between the relative merits of reductionism and holism (Koestler 1969, Edwards and Jaros 1994, Naveh 2000) and turns out to be particularly adapted for complex systems modelling, as demonstrated by many studies (Kay and others 1999, Giampietro and others 2006, Kira and van Eijnatten 2008).

Given the holistic framework imposed by INRM, the genericity of the modelling language is a prerequisite in this context. Examples of generic languages are the data modelling language EXPRESS

or the Unified Modelling Language (UML). Nevertheless, when dealing with multiple fields of knowledge, generic languages have a substantial drawback related to semantic issues, as mixing terminologies of various domains usually results in more ambiguity. In contrast, domain-specific languages are specifically created to deal with problems of a particular domain and are not intended to be used outside their predefined environment. On the other side, domain-specific modelling languages may overcome the semantic issues linked to genericity and therefore facilitate the use by modellers, as shown by Fall (2001) with the SELES language developed for landscape modelling.

### Ontology

Ontology languages are used to formalize and structure knowledge domains, and therefore allow to clarify semantic issues associated with the integration of multiple knowledge sources. A review of the various ontology languages has been carried out by Corcho (2003), whereas until now new developments and new languages have been meanwhile the subject of an abundant literature. Examples of such languages are the Web Ontology Language (OWL), IDEF5 or KIF. In information science, the term “ontology” refers to a set of concepts and relationships between these concepts used for describing a field of knowledge. One of the major contributions towards the foundations of an ontological framework has been carried out by Bunge and extended by Wand and Weber (Wand 1999). Known as the Bunge-Wand-Weber (BWW) model, this framework proposes “a set of high-level, abstract constructs that are intended to be a means of representing all real-world phenomena”. These ontological constructs are often used as references to evaluate the grammar constructs of any modelling language (Opdahl and Henderson-Sellers 2001, Opdahl and Henderson-Sellers 2004).

Having a modelling language that handles ontological information makes the modelling of any kind of system effective, provided a consistent ontology of the field exists beforehand. In other words, the ontology allows users to agree on a common “vocabulary” required for the system description. Several examples of such ontologies related to the field of environmental management may be found in the literature. For instance the OntoWEDSS that uses an ontology of microbiologic knowledge to model the wastewater treatment process (Ceccaroni and others 2004), the “system model for water management” of Schenk and others that provides an ontology of the water system and its related elements to support the process of integrated water resources management (Schenk and others 2009) or the Extensible Observation Ontology (OBOE), a formal ontology that aims to “capture the semantics of generic scientific observation and measurement” (Madin 2007).

### Modelling systems

Systems modelling languages are generally used in computer science, systems engineering or project management. Families of languages such as UML or IDEF can cover a wide range of uses, principally in order to follow all the steps of the software development lifecycle. Other languages such as SysML (Willard 2007), Modelica (Mattsson and others 1998) or USL (Hamilton and Hackler 2008) are specially adapted for advanced systems engineering. It is interesting to note that these few examples are only a sample of the many languages available in these fields, whereas languages dedicated to environmental modelling are very rare.

One of the rare examples related to environmental modelling is the Energy System Language (ESL) developed by Howard T. Odum (1960) which allows modelling the ecosystems through energy flow diagrams. ESL is based on the analogies between energy flows within the ecosystems and those in electronic circuits and has been applied successfully in a variety of situations from ecological to economic modelling (Odum and Odum 2000, Brown 2004, Rivera and others 2007). Although it is a powerful language, ESL remains too specific regarding other kind of systems models such as causal loop diagrams or Bayesian networks.

Finally, in the frame of systems dynamics, which involve the use of causal loop and stock and flow diagrams, some approaches such as Simile (Muetzelfeldt and Massheder 2003) and STELLA (Costanza and Gottlieb 1998) incorporate their own graphical modelling language. These languages remain unfortunately specific to their modelling task, and therefore difficult to apply in other contexts.

### Modelling interactions

Modelling interactions has been identified as the critical issue concerning both systems modelling and ontology languages available in the literature. On the one hand, systems languages (such as UML, Petri-Net, SysML, ESL...) have interactions only dedicated to a single specific task, such as stating that a system attribute has a particular impact on another system attribute or quantifying a flow between two elements. For instance, in the framework of systems dynamics, the influences within causal-loop diagrams may be interpreted as simplifications of complex phenomena. Taking the case of the “predator-prey model” that is used to describe the joint evolution of two animal species, one may consider dissecting the simple influences between the animals’ populations, death rates and birth rates, in order to represent with more details the systems’ complexity. The positive influence of the prey’s abundance on the predator’s birth rate could be represented by a set of

systems and interactions that take into account the flows of “nutritive elements” supplied by the prey, the geographic locations of the prey and predators, etc.

On the other hand, ontology languages don’t provide a better framework for modelling interactions. The latter may be defined through properties or meta-classes. At first, simple interactions may be described through properties, the “domain” containing the possible sources and the “range” embracing the targets. But this approach doesn’t allow detailing interactions. An alternative is the use of a meta-class (i.e. a class whose instances are classes) to define the interactions, as shown by the “engineering ontologies” provided in documentation of the ontology editor Protégé (Tudorache 2008). For instance, a flow of phosphate-contaminated water could be derived from a class with four properties: “hasSource”, “hasTarget”, “hasValue” and “hasPhosphateConcentration”. However, ontology languages being often restricted to the first order logic, only few ontology languages (among which OWL-Full) propose the use of meta-classes.

Although interactions may be modelled with conventional ontology languages, there are several advantages to adopt a language that integrates interaction as a basic construct. At first, it will put the interactions at the same conceptual level than the other systems elements, be they objects or information. Then, considering graphical languages, it offers the possibility to assign an appropriate symbol to interactions and therefore allows creating diagrams where interactions are illustrated by arrows, instead of being represented by “boxes” like other regular system elements.

### Graphical vs. textual language

Compared to a textual or a mark-up language, a graphical language may be easier to understand by non-specialists. The diagrams produced by graphical languages seem particularly adapted for describing systems, as they are, by definition, composed of a set of interconnected elements. Nevertheless, the difference between graphical and textual language isn’t essential. One can imagine to translate a graphical language into a mark-up format (such as xml) and inversely, like GrOWL (Krivov and others 2007) is a visualization of the OWL mark-up language.

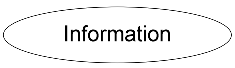
# The SysMod language

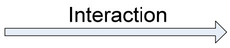
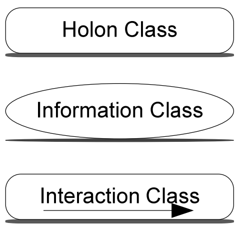
The SysMod (System Modelling) language has been developed to take advantage of the particularities of ontology and systems modelling languages and to consider in depth the modelling of systems interactions.

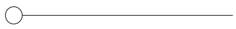
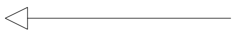
### Language constructs

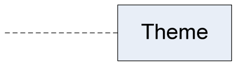
The constructs are the high-level syntactic elements of the modelling language. Table 1 summarises the constructs that compose the SysMod language, along with the symbols that have been proposed to represent them graphically.

Table 1: The SysMod list of constructs

| Constructs | Explanations |
|---|---|
| System element | |
| Holon  | In the SysMod language, a holon is almost equivalent to the "thing" of the BWW ontological model (Opdahl and Henderson-Sellers 2004) or to the "individual thing" of Chisholm's ontology (Chisholm 1996). A holon has the particularity to be simultaneously a whole and a part, i.e. a system in itself and a part of a larger system. Holons are used for modelling the real world as a set of hierarchically embedded systems. |
| Information  | Information symbols represent any kind of data that can describe or document a system. It can be a text, a spatial geometry, a numeric value, a fuzzy value, a file, an image, etc. |
| Property<br><br> | Like BWW things, holons possess properties. However, the SysMod definition differs from the conventional meaning in the sense that SysMod properties, when they consist in groups of holons or groups of information, only act as kinds of repositories for the real instances. The latter are connected to the property through the membership relationship. SysMod property is the generalization of three types of system elements<br><br>As its name suggests, a group of holons gathers a number of holons. For instance, a bicycle may have a property "wheels of the bicycle", which is a group that has a cardinality of two. |

| | |
|---|---|
| <br> | In the same way, a group of information can host a set of data. For instance, the bicycle may have a speed property that groups speed values for a time period.<br><br>The interaction is a special kind of property, in the sense that it is at the same time a system element and a relationship. It is a relationship because it can connect groups together and it is a system element because it can carry other elements, for instance a flow of material or a flow of information. An interaction can only be connected to groups. If the interaction is a flow, its source and target are elements that act as (sub)-system input, output or stock; the latter can only be modelled through the use of groups. Secondly, when the interaction is an influence, there is a SysMod convention establishing that the interaction needs some properties (in fact groups) as source and target. For instance, in the case of a hydrological model, the influence of a region on the water runoff will be modelled through an influence between the property “infiltration rate” of the region and the property “runoff” of the same region.<br>When the existence of an element acts on others, this impact isn’t represented by an interaction, but is modelled through a “property role” relationship whose type is “Acts on”. |
| Class<br> | A class is a definition, a template, for system elements sharing one or more properties. One can consider the class as a blueprint for the creation of instances. SysMod distinguishes three types of classes: the holon class, the information class and the interaction class. Each of them gathers the instances of the corresponding type. |
| Relationships | |
| Property role<br>- (not defined)<br><br>Roles defining the holon’s internal components<br>- Contains<br>- Consumes<br>- Produces<br><br>Roles describing the relationships between the holon and its immediate | The “property role” relationship is used to connect properties with the system elements they characterize.<br><br>In the case of properties being internal components of a holon, one can distinguish three types of “property roles”: contain, consume and produce. In the case of modelling a reservoir, one can for instance assign three properties to the instance of reservoir: water stock, water input and water output, these properties having respectively the three different aforementioned roles.<br><br>Three additional “property roles” are added in order to consider the aptitude of holons to interact with their environment. These roles |

| | |
|---|---|
| environment<br>- Possesses<br>- Acquires<br>- Releases<br><br>Roles of the holon as a part of a bigger system<br>- Acts on / manages<br>- Constructs<br>- Destroys | allow the holon to acquire, possess and release some elements. In the context of the reservoir example, one could say that the reservoir "possesses" a source of water supply. These "roles" may also be used to characterize activities of living being such as a person "acquires" a car and "releases" some wastes.<br><br>The last type of "property roles" is used for modelling the impact of the existence of a holon on the other systems elements. These roles are particularly useful to determine the function of a specific holon as a part of a bigger system. It can be the role of a piece in a machine, the role of an organ in a living body or the role of a person within an organization. |
| General instantiation<br> | The general instantiation relationship is used for creating a property from a class. For instance, the "weight of Mr X" is a property (in this specific case a group of information) that is instantiated from the "weight" class. |
| Particular instantiation<br> | The particular instantiation relationship connects an instance to its corresponding class, for instance "80 kg" to the "weight" class. |
| Membership<br> | The membership relationship is used for linking a "group of holons" or a "group of information" to its members (respectively holons and information elements). Taking the example of the "weight of Mr X", the instance members of the group could be 80 kg at one time and 83 kg at another time. |
| Generalization/Specialization<br> | The generalization/specialization relationship represents a link between two classes where one class (the specialization) is a subclass of the other one (the generalization). This allows modellers to create a taxonomy of classes, a first step towards the realization of an ontology. |
| Mereotopology<br> | Mereotopology combines both set of relationships from mereology (relationships dealing with parts and their respective wholes) and topology (relationships about the spatial properties of elements, such as connectedness). Several mereotopological theories exist, such as GEMTC (Varzi 1996) or RCC (Cohn and others 1997, Bennett and others 2002).<br>The theory used by SysMod may be changed or extended according to the modelling requirements. As an illustration, the axioms of the theory RCC-8 are listed below (Cohn and others 1997):<br>1. DC(x; y) is disconnected from y<br>2. EQ(x; y) x is identical with y<br>3. PO(x; y) x partially overlaps y<br>4. EC(x; y) x is externally connected to y |

| | |
|---|---|
| | 5. TPP(x; y) x is a tangential proper part of y<br>6. TPPi(x; y) y is a tangential proper part of x<br>7. NTPP(x; y) x is a nontangential proper part of y<br>8. NTPPi(x; y) y is a nontangential proper part of x |
| Other constructs | |
| Theme<br><br> | SysMod themes consist of sets of system elements. The resulting thematic groups allow the modeller to (1) create scenarios, (2) characterize events, (3) develop strategies, (4) create indicators sets or (5) make arbitrary groups of system elements. |
| (1) Scenario | Multiple potential futures of a system can be differentiated through the use of scenarios. For instance, two distinct scenarios for a city could be a high demographic growth or a stabilization of the population, each of them having different values for the property population. |
| (2) Event | Events are created by selecting sets of involved system elements for a certain time slot. A flood event could be described, for instance, with some values series about precipitations, river flows and runoffs for a given system, along with some spatial information on the extent of the flood. |
| (3) Strategy | Strategies are defined to describe the answer of a "system manager" to a given scenario, through, for instance, a modification of the system structure or composition. |
| (4) Indicators set | Obtained by bringing into a same theme some system properties, the indicators sets may be useful for targeting or monitoring purpose. For example, such indicators set could be the list of the Millennium Development Goals for a specific country. |
| (5) Arbitrary set of system elements | A theme can also be used for tagging arbitrarily a set of system elements. |
| Time context<br>- Start (yyyymmddThh:mmZ)<br>- End (yyyymmddThh:mmZ)<br>- Period (yyyymmddThh:mm) | Any system element can be marked with a "time context". This construct includes a time start and a time period whose format is based on the Standard ISO 8601.<br>Given the fact that multiple scenarios may involve different lifetime for a system element, the latter can therefore host several "time contexts", i.e. one per connected theme. |

## SysMod meta-model

To improve the understanding of the SysMod constructs and to promote possible comparisons with other languages, a meta-model of SysMod has been drawn (figure 1). This meta-model shows the main classes of concepts of the language and the relationships between them. It is expressed as an entity-relationship (ER) model and can therefore be translated directly into a relational database. An example of such development has been carried out by Schenk and others (2010a), who created a database for environmental information based upon the SysMod language.

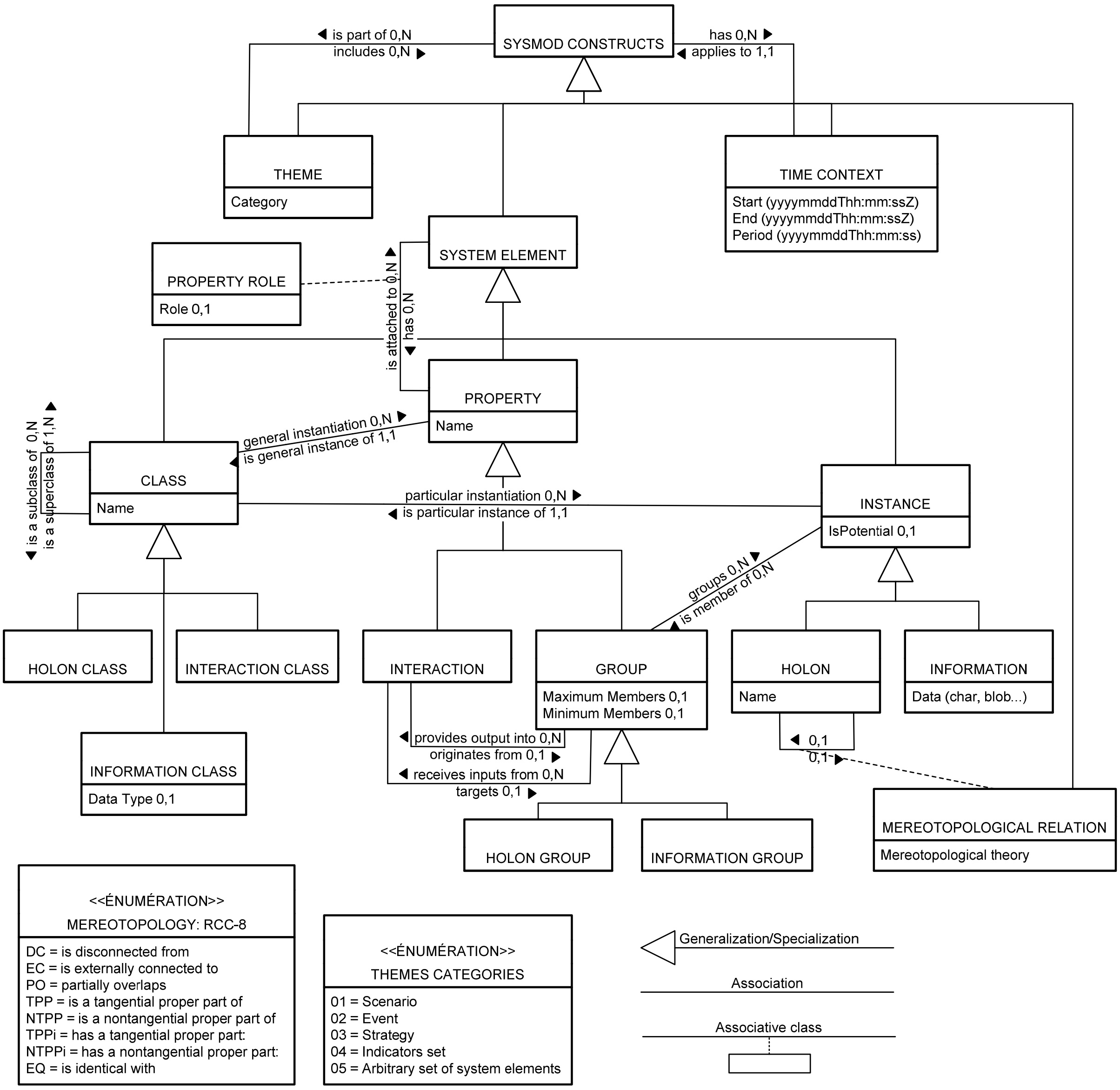



Figure 1: SysMod meta-model

A similar process has been realised by Rosemann (2004), who converted the BWW model constructs into an ER-based meta model containing 50 entity types and 92 relationship types, in order to better explain core constructs of the BWW ontology.

Moreover, the use of relational database to host systems data presents an alternative to the conventional ontology languages which rely on the use of mark-up languages, like for instance the Web Ontology Language OWL which is based on XML syntax. Moreover, relational database management systems such as PostgreSQL also offer proven tools and techniques to control the creation, the maintenance and the use of data and therefore may be a valuable contribution towards system modelling.

# Modelling examples

### Ontologies

As noted by Villa, "ontologies can be used at different degrees of internal complexity and expressive power" (Villa and others 2009). At a basic level, ontologies consist in simple inventories of concepts that are not or loosely connected. Examples of such ontologies are basic thesauri or vocabulary indices. At a more complex degree, taxonomies are more organized ontologies, composing hierarchies of classes made through the use of generalization/specialization relationships. Figure 2 illustrates the case of three taxonomies: a taxonomy of holons that covers the animal reign, a taxonomy of information types and a taxonomy of interaction types. Finally, at the highest degree of complexity, detailed ontologies are obtained when classes, properties and possibly instances are combined, such as illustrated by the Lynx's ontology (figure 3).

### Systems components and systems dynamics

An important feature of the SysMod language is its ability to create instances and groups of instances from the classes defined in ontologies. This allows the modeller to build a model using class instantiations.

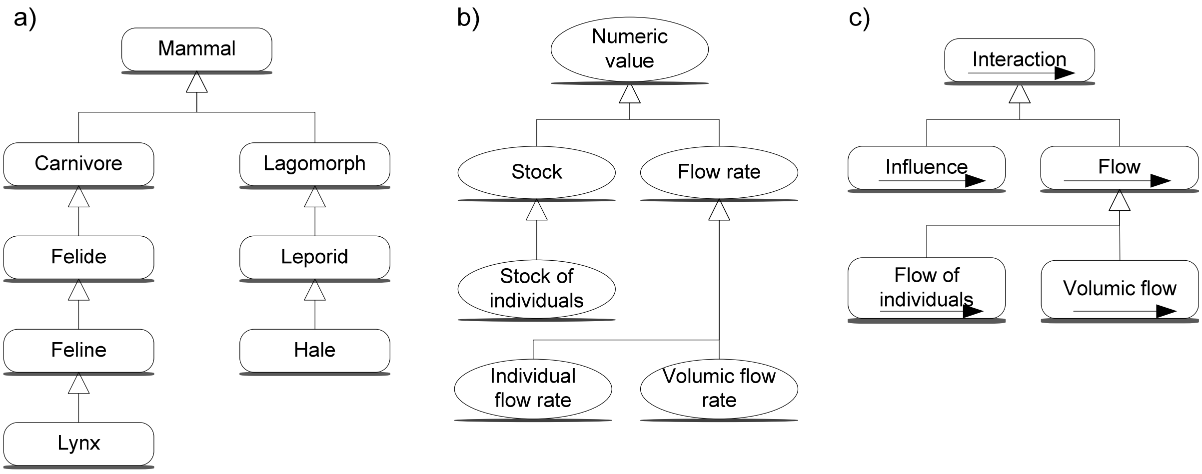



Figure 2: Examples of basic taxonomies of classes: (a) ontology of holon classes, (b) ontology of information types and (c) ontology of interaction types

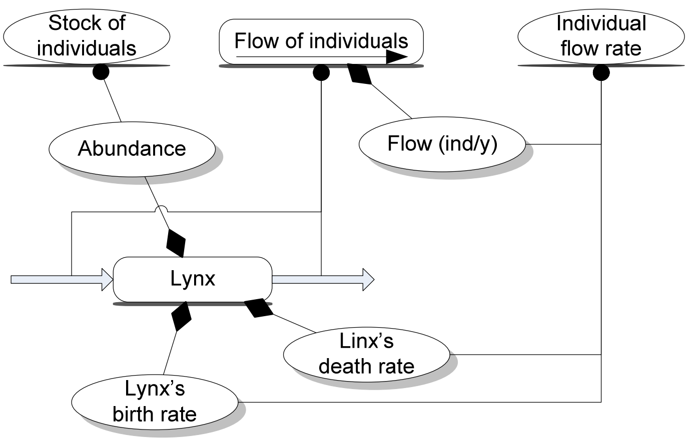



Figure 3: More detailed ontology that combines the various SysMod constructs

Based on the elements of the traditional predator-prey system, figure 4 illustrates the modelling of a simple system composed by two populations of animals (the predators and the preys) and their respective properties. This system can be extended to model the different influences between properties, such as in traditional stock and flow diagrams, except that the graphical symbolism is different (figure 5).

As highlighted in the previous section, interactions should be considered as complex systems elements. Thus, a flow may be characterized by several properties, as for instance a stream of polluted water has properties “flow rate” and “pollutant concentration”. There is the case where an

interaction summarizes a complex system behaviour in a single relationship. This situation is illustrated in figure 6, where the influence of the lynx's abundance on the hare's death corresponds to a composition of many other interactions.

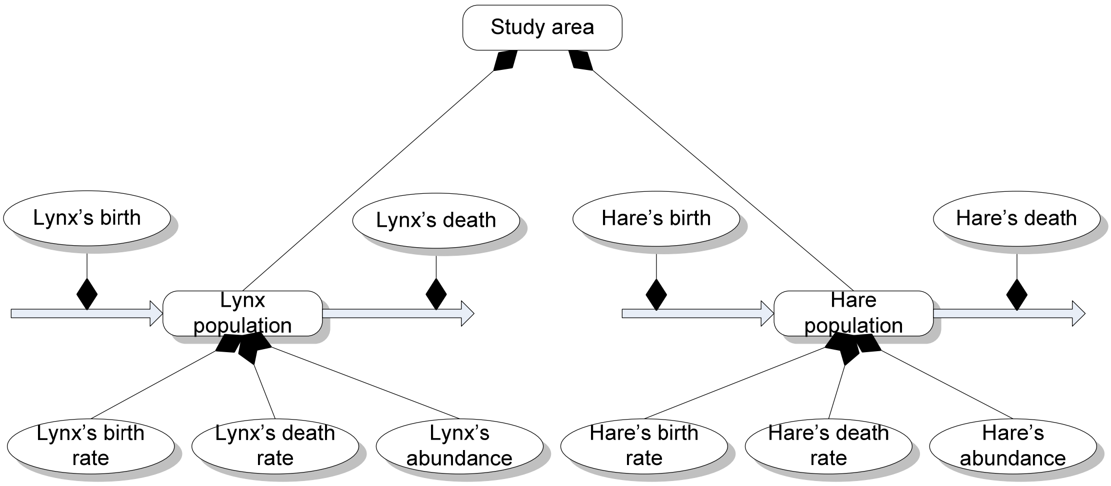



Figure 4: Modelling of system components, the case of the predator-prey model

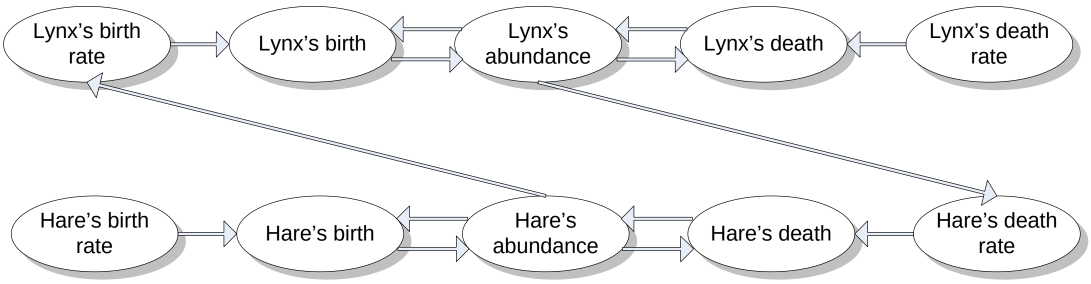



Figure 5: Modelling of system interactions, the case of the predator-prey model

## System definition and boundaries

As mentioned before, the systems of holons have the particularity to be organized hierarchically. Consequently, they can be examined from different angles. At first, the holon can be perceived as a whole, i.e. a system in itself with its own identity. This is shown by the class "Human being" and the two instances of persons in figure 7, each of the three systems being composed by subsystems. This figure also highlights the difference between what can be called the functions and the structure of the system, as described by Giampetro (2006) in his article on "the epistemiological predicament". On the one hand, the "heart" property represents a "functional entity", given that it corresponds to

an internal function of the system that allows him to work (or stay alive). On the other hand, the instances "human heart" and "artificial heart" are two different "structural entities", although they perform the same functions (i.e. they correspond to an equivalent property "heart"). One may also notice in this figure the issue concerning the definition of the hierarchical levels: e.g. does the "heart" is a property of the "human body", or is it the property of an intermediate holon such as the circulatory system?

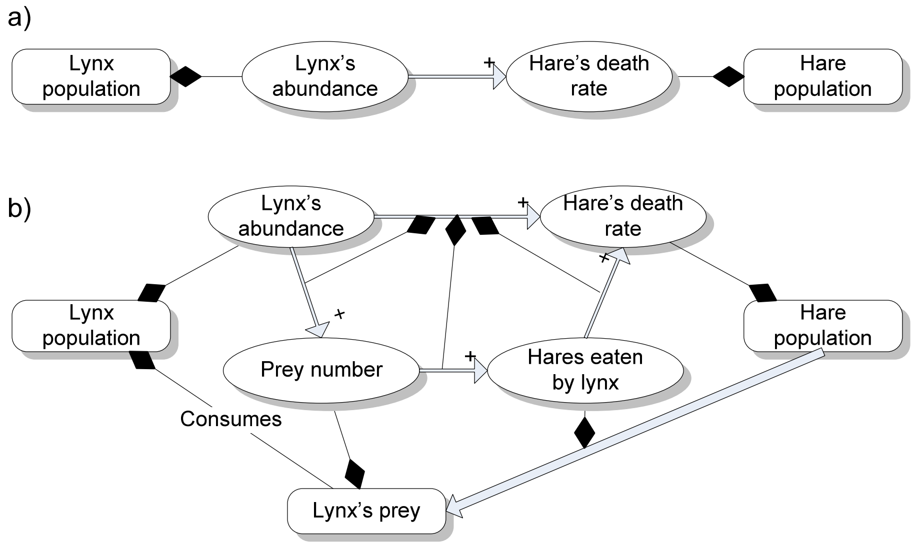



Figure 6: Interactions as complex systems elements: (a) a simple influence may be decomposed in (b) several parts

Secondly, the holon may be considered as a system interacting with its environment. This is particularly important when the holon is a living being, and therefore has the capability to act on its environment. The next example shows the use of different "property roles" (possesses, gets, parts with) to illustrate the microsystem that surrounds the holon (figure 8).

In the third case, the holon may be a part of a larger system, as illustrated by the figure 9, where Mr X is an instance with the function of director within a larger system, the Enterprise Z.

### Time, spatiality and multiple potential scenarios

The last example demonstrates the use of thematic clustering to represent specific systems characteristics such as events, scenarios, strategies or indicators' sets. Indeed, themes allow the modeller to isolate some sets of system elements in order to describe special system features such as

events or scenarios. Figure 10 illustrates the use of "thematic tags" to characterize events for different scenarios.

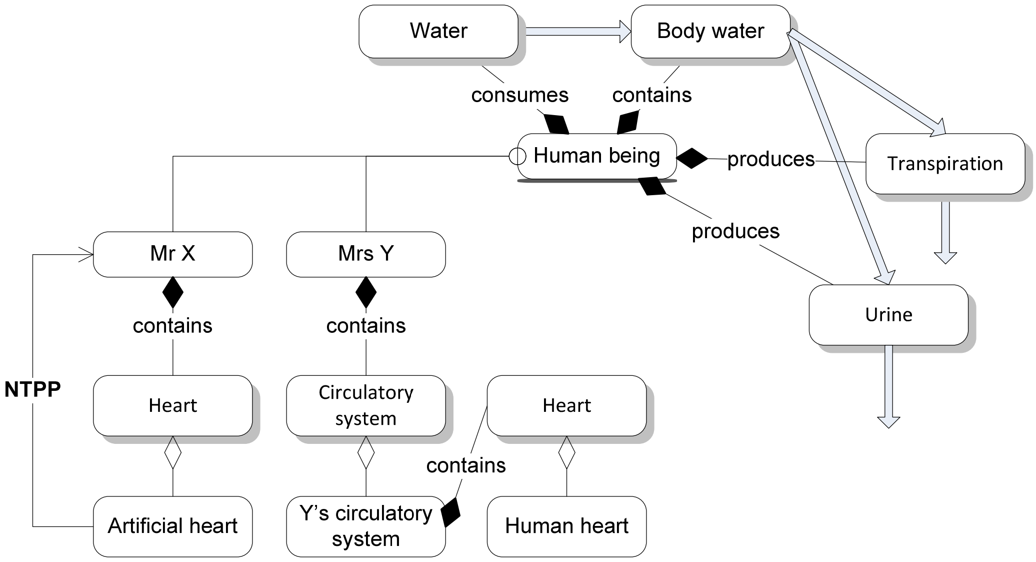



Figure 7: The internal composition of holons, (on the right) example of the water flow within the human body and (on the left) example of two different structural compositions with the same function. The mereotopological relationship NTPP means that the artificial heart "is a nontangential proper part of" Mr X.

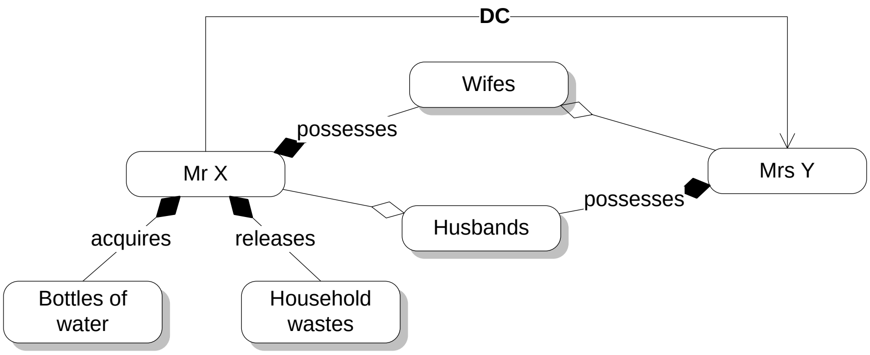



Figure 8: The direct environment of a holon forms a kind of microsystem which is an external characterization of itself. The relationship DC means that Mr X is topologically "disconnected from" Mrs Y.

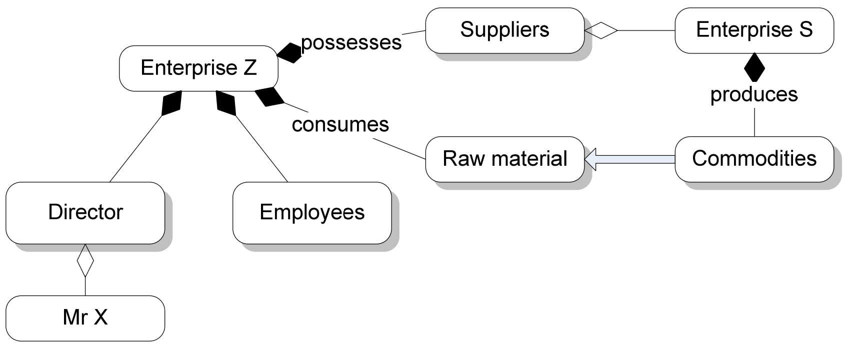



Figure 9: The holon (Mr X) as a part of a bigger system (the holon Enterprise Z)

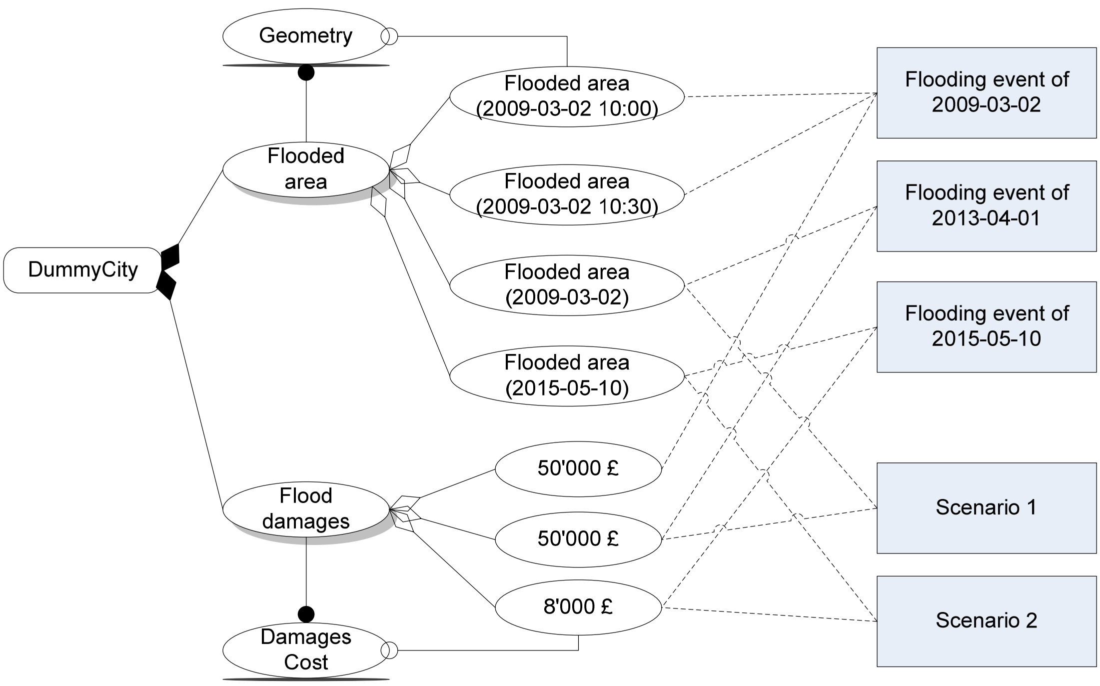



Figure 10: Use of themes to describe events and scenarios

# Discussion

The SysMod features were illustrated by some various examples that demonstrated the aptitude of the language to model both ontologies and systems within a common framework. Furthermore,

using the concept of holon, although this notion is neither widespread in the field of ontology nor in system modelling, it highlights the hierarchical architecture of systems.

However, this duality between ontology and system model presents two major drawbacks. The first is the dependence of systems modelling on ontologies; i.e. the development of an ontology is a prerequisite to model any system. This creates extra work for the modellers, compared to ready-made domain-specific languages. The second issue concerns the limitations of SysMod in the field of ontology engineering. SysMod doesn't have indeed the linguistic wealth of other ontology languages, which may use properties that express elaborate relationships between systems elements. For instance, according to the OWL wine ontology developed by World Wide Web Consortium (Smith and others 2004), the property "madeFromGrap" characterizes the relationship between the domain "Wine" and the range "WineGrape". The property summarizes in three words the process of transformation from grapes to wine. In comparison, SysMod would require to model the whole winemaking process to express this link. However, this is not really a shortcoming since this restriction prevents modellers from being fooled by linguistic shortcuts and obliges them to really consider their model as an operational system.

By placing these developments in the context of the integrated natural resources management (INRM), the use of tools based on SysMod language would require to take into account the following steps:

1. Define the goals of the systems modelling and investigate the different fields of knowledge that should be integrated. This could be the result of a participative process that involves stakeholders and modellers.
2. Create an ontology that covers the different classes required for the system description, ideally a kind of glossary (organized in a taxonomic or ontological way) that emerges from consensus between users.
3. Model the system through the instantiations of these different classes. In order to take into account the different perceptions of the stakeholders, several models may be created and then compared.

4. Through systems analysis or simulations, test and validate the system model with data from the real world.
5. If required, modify the system model or the ontology in order to get results that match the modelling goals.

In its actual state, SysMod has been used with the support of a PostgreSQL database and then takes advantage of the functionalities offered by the database management system (DBMS) as regards the creation and maintenance of large amount of data. This situation involves that the structure of the data (defined by the ontology) and the data themselves are stored in the same location. It seems particularly convenient as the users don't have to understand and modify the database structure, but it may also present a risk as a slight change in the ontology may have severe repercussions on the data validity.

# Conclusion and outlook

The present paper proposes a new modelling language – SysMod –designed firstly for modelling environmental systems, but susceptible to be applied to any kind of systems. SysMod is a generic, graphical language whose main particularity is the ability to create ontologies, in order to ensure the semantic coherence of models and to provide predefined elements for systems creation. SysMod also distinguishes itself from other modelling languages through its ability to consider systems interactions as complex systems elements.

To deal with complicated systems, some software developments may be very useful to create and manage models using SysMod, such as a diagramming tool that allows modellers to create and dynamically explore systems (Schenk and others 2010b), a geographic viewer to display spatial attributes or a reporting solution (Schenk and others 2010b). In the future SysMod may give rise to modelling tools which could compute the systemic models and make use of external models to run simulations.

# References

Bellamy JA, Walker DH, McDonald GT, Syme GJ (2001) A systems approach to the evaluation of natural resource management initiatives. Journal of Environmental Management 63 (4): 407--423.

Bennett B, Cohn AG, Wolter F, Zakharyaschev M (2002) Multi-Dimensional Modal Logic as a Framework for Spatio-Temporal Reasoning. Applied Intelligence 17 (3): 239-251.

Brown MT (2004) A picture is worth a thousand words: energy systems language and simulation. Ecological Modelling 178 (1-2): 83-100.

Ceccaroni L, Cortés U, Sànchez-Marrè M (2004) OntoWEDSS: augmenting environmental decision-support systems with ontologies. Environmental Modelling & Software 19 (9): 785-797.

Chisholm RM (1996) A Realistic Theory of Categories: An Essay on Ontology Cambridge University Press, New York, 146 pp.

Cohn AG, Bennett B, Gooday J, Gotts NM (1997) Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. GeoInformatica 1 (3): 275-316.

Corcho O, Fernández-López M, Gómez-Pérez A (2003) Methodologies, tools and languages for building ontologies. Where is their meeting point? Data & Knowledge Engineering 46 (1): 41-64.

Costanza R, Gottlieb S (1998) Modelling ecological and economic systems with STELLA: Part II. Ecological Modelling 112 (2-3): 81-84.

Edwards LB, Jaros GG (1994) Process-based systems thinking--Challenging the boundaries of structure. Journal of Social and Evolutionary Systems 17 (3): 339-353.

Fall A, Fall J (2001) A domain-specific language for models of landscape dynamics. Ecological Modelling 141 (1-3): 1-18.

Giampietro M, Allen TFH, Mayumi K (2006) The epistemological predicament associated with purposive quantitative analysis. Ecological Complexity 3 (4): 307-327.

Hamilton MH, Hackler WR (2008) Universal Systems Language: Lessons Learned from Apollo. Computer 41 (12): 34-+.

ICWE (1992) The Dublin Statement on Water and Sustainable Development. Accessed online January 9: http://www.wmo.int/pages/prog/hwrp/documents/english/icwedece.html.

Kay JJ, Regier HA, Boyle M, Francis G (1999) An ecosystem approach for sustainability: addressing the challenge of complexity. Futures 31 (7): 721-742.

Kira M, van Eijnatten F (2008) Socially Sustainable Work Organizations: A Chaordic Systems Approach. Systems Research and Behavioral Science 25 (6): 743-756.

Koestler A (1967) The ghost in the machine. Hutchinson, London :.

Koestler A (1969) Beyond atomism and holism — the concept of the holon. In: Koestler A and Smythies JR (ed) Beyond reductionism; new perspectives in the life sciences, New York, pp 192–216.

Kogut P, Cranefield S, Hart L, Dutra M, Baclawski K, Kokar M (2002) UML for ontology development. The Knowledge Engineering Review 17 (1): 61-64.

Krivov S, Williams R, Villa F (2007) GrOWL: A tool for visualization and editing of OWL ontologies. Web Semantics: Science, Services and Agents on the World Wide Web 5 (2): 54-57.

Madin J (2007) An ontology for describing and synthesizing ecological observation data. Ecological informatics 2 (3 SPEC. ISS.): 279-296.

Mattsson SE, Elmqvist H, Otter M (1998) Physical system modeling with Modelica. Control Engineering Practice 6 (4): 501-510.

Midgley G (1992) Pluralism and the legitimation of systems science. Systemic Practice and Action Research 5 (2): 147-172.

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

Mulej M, Potocan V, Zenko Z, Kajzer S, Ursic D, Knez-Riedl J (2004) How to restore Bertalanffian systems thinking. Kybernetes 33 (1): 48-61.

Naveh Z (2000) What is holistic landscape ecology? A conceptual introduction. Landscape and Urban Planning 50 (1-3): 7-26.

Odum HT (1960) ECOLOGICAL POTENTIAL AND ANALOGUE CIRCUITS FOR THE ECOSYSTEM. American Scientist 48 (1): 1-8.

Odum HT, Odum EC (2000) Modeling for all scales: an introduction to system simulation. San Diego, California, USA, 80 pp.

OMG OMG (2009) Ontology Definition Metamodel. In.

Opdahl AL, Henderson-Sellers B (2001) Grounding the OML metamodel in ontology. Journal of Systems and Software 57 (2): 119-143.

Opdahl AL, Henderson-Sellers B (2004) A template for defining enterprise modelling constructs. Journal of Database Management 15 (2): 39-73.

Rammel C, Stagl S, Wilfing H (2007) Managing complex adaptive systems - A co-evolutionary perspective on natural resource management. Ecological Economics 63 (1): 9-21.

Rivera EC, de Queiroz JF, Ferraz JM, Ortega E (2007) Systems models to evaluate eutrophication in the Broa Reservoir, São Carlos, Brazil. Ecological Modelling 202 (3-4): 518-526.

Rosemann M, Green P, Indulska M (2004) A Reference Methodology for Conducting Ontological Analyses. In: (ed) Conceptual Modeling – ER 2004, pp 110-121.

Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A (2010a) A systems-based, generic environmental data model. Environmental Management Submitted.

Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management. Environmental Management 43 (3): 458-469.

Schenk C, Roquier B, Soutter M, Mermoud A (2010b) A generic decision support system. In preparation.

Smith MK, Welty C, McGuinness DL (2004) OWL Web Ontology Language Guide. Accessed online January 22, 2010: http://www.w3.org/TR/owl-guide/.

Tudorache T (2006) Employing Ontologies for an Improved Development Process in Collaborative Engineering. In *Institut für Telekommunikationssysteme* Technische Universität Berlin, Vol. Dipl.–Ing, p 171.

Tudorache T (2008) http://protegewiki.stanford.edu/index.php/Engineering_ontologies. Accessed online January 21, 2010: http://protegewiki.stanford.edu/index.php/Engineering_ontologies.

Varzi AC (1996) Parts, wholes, and part-whole relations: The prospects of mereotopology. Data & Knowledge Engineering 20 (3): 259-286.

Villa F, Athanasiadis IN, Rizzoli AE (2009) Modelling with knowledge: A review of emerging semantic approaches to environmental modelling. Environmental Modelling & Software 24 (5): 577-587.

Wand Y (1999) An ontological analysis of the relationship construct in conceptual modeling ACM TRANSACTIONS ON DATABASE SYSTEMS. Acm Transactions on Database Systems 24 (4): 494-528.

Willard B (2007) UML for systems engineering. Computer Standards & Interfaces 29 (1): 69-81.

# ANNEX II

## ART: AN ACTIVE SYSTEMS-BASED NAVIGATION AND REPORTING TOOL

# Preface

This annex presents a reporting tool which allows displaying and managing data related to system elements in lists, as well as in dynamic, user-friendly report sheets.

This annex is adapted from the article: Soutter M, Roquier B, Schenk C, Mermoud A (2010) ART: An active systems-based navigation and reporting tool, submitted to Information Systems Frontiers.

# Abstract

Information systems, decision support systems or integrated modeling frameworks were given a lot of attention in recent years, since these tools are to facilitate the spreading and generalization of the global, integrated and holistic approaches advocated as a necessary step to more sustainable decision making.

This paper presents a dynamic reporting tool that forms part of a more general information and knowledge sharing platform (called the generic decision support system GenDSS). This reporting tool provides the framework requested to elaborate rich template descriptions that will give added value to the linked database items they refer to.

Thanks to the systemic nature of the database backing the reporting tool, several additional novel features are provided, that turn this reporting tool into a database navigation tool. As such it complements perfectly the systemic and geographic navigation tools already present in the GenDSS platform.

# Introduction

Addressing issues in a global, integrated, holistic way - as opposed to the sectorial, piecemeal approaches that are still in widespread use in many areas of environmental management - has long been and is still advocated as an inescapable step on the way towards more sustainable decision making processes. This has been the case for instance in water resources management, where concerns and decisions are usually split among a set of rather weakly linked stakeholders, taking care in an almost autarkic way of their own specialized area: drink water supply, waste water disposal, flood control, aquatic ecosystems, energy production, recreational use of water bodies, and so on.

Environmental planning and management problems are in most cases complex, as they encompass many different areas of knowledge and science, and involve a lot of different stakeholders, each of them with its own concerns, background, perspectives, agenda and objectives. It is therefore not surprising that data accessibility, data sharing and integration as well as knowledge and models sharing and integration, have become major issues in contemporary environmental research. A lot of effort has been made in recent years to develop a whole range of tools that facilitate this integration, including information systems (IS), decision support systems (DSS) or integrated modelling frameworks (IMF).

Data sharing is much a matter of data retrieving and presentation. Classically, these techniques include online analytical processing methods (OLAP), reporting, and data mining (O'Brien and Marakas 2009). OLAP focuses on multi-dimensional queries and generally present the results in the form of matrices (Pedersen and Jensen 2001). Data mining is mostly used to uncover patterns in datasets (e.g. spatial continuity, time dependencies) (Kantardzic 2003). Whereas both techniques provide analytical tools, reporting is centred on more static data presentation layouts; it is less concerned by data exploration than by communication to possibly numerous users (Shadan 2005). The reporting process encompasses data extraction from a data warehouse, and possibly its subsequent transformation and load, and presentation with a reporting tool. These tools, such as the open source BIRT (Eclipse Foundation 2009) or JasperReports (Jaspersoft 2000), and the commercial ActiveReports (GrapeCity 1997) or CrystalReports (SAP 2009), to mention just a few, may achieve the whole process or be limited to the last presentation step. They however suffer from a certain lack of flexibility regarding their integration into a broader enterprise reporting system or a decision support

system, since changes in the source code are usually needed to provide the appropriate report management facilities (Ni and others 2007). Embedding a reporting tool in a broader DSS, as is the case for the hereby presented ART, gives added-value to the tool itself: linking the reporting tool with other data management and presentation tools (following spatial and/or system logics in our case), along with synchronization mechanisms, enhances and emphasizes the navigation and exploration functionalities of a reporting tool.

The generic decision support system GenDSS (Schenk and others 2010a) ART belongs to, is developed within the frame of the EU SWITCH project (Sustainable Water Improves Tomorrows Cities' Health). This generic, versatile, system-based, decision support system has been setup and implemented for the purpose of improving actual integrated urban water management in the project's partner cities. This GenDSS was primarily designed as a knowledge and information sharing platform to support local Learning Alliances' activities. As a consequence, this tool follows a twofold development line: on the one hand it is meant to let users access and share easily a wide range of information, i.e. it is a communication tool with a quite large group of non specialized end-users, and on the other hand, it is meant as a more targeted tool allowing users to explore various alternatives for the future, in order to address issues in current decision-making related to water infrastructures.

As a communication platform, the GenDSS (figure 1) provides, among others, a set of complementary tools to browse information in different ways: (i) a system viewer, that takes advantage of the system-based structure of the database (Schenk and others 2010b) to navigate the information along its system logics, (ii) a geographic viewer, that allows to navigate information on the basis of its spatial position and (iii) an active or dynamic (since dynamically updated with changes in the linked database) browser-like reporting tool, that on one hand provides direct access to the raw data stored in the database, and on the other hand allows for contextualization of those raw data, within report sheets. The present paper aims at detailing somewhat the structure, layout and features of this reporting tool.

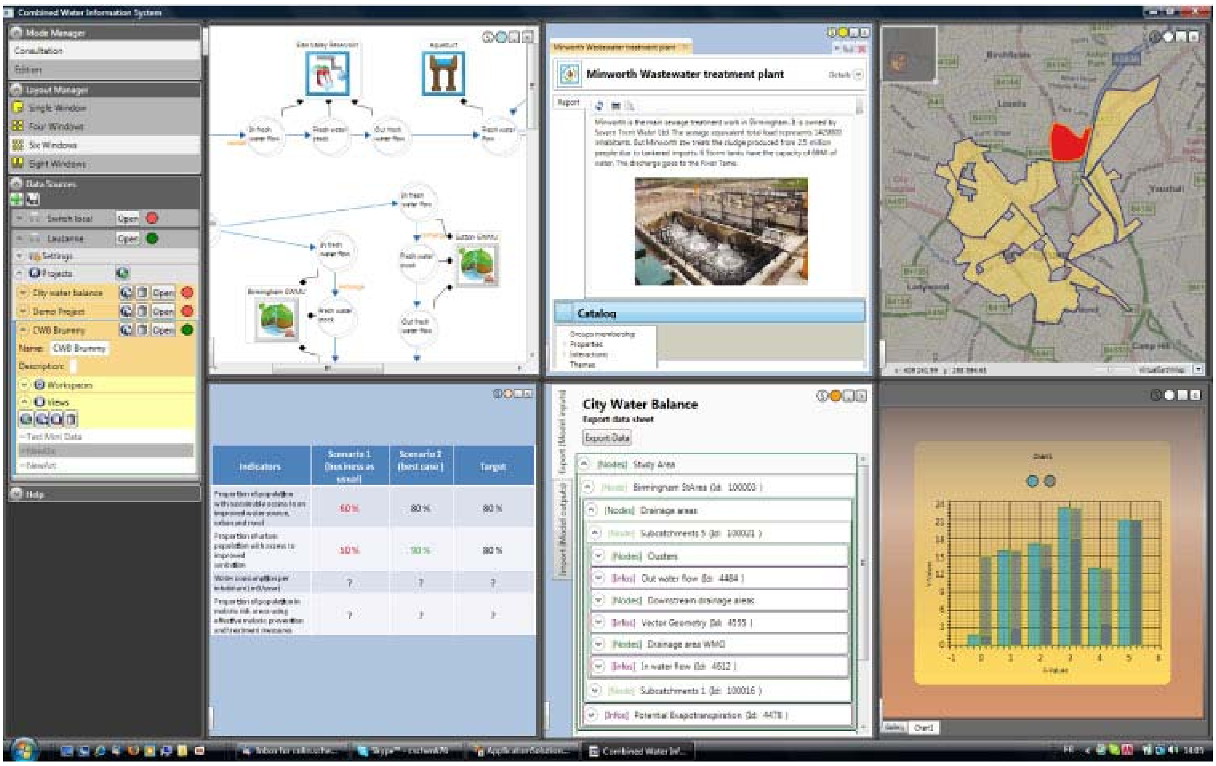



Figure 1: Overall GenDSS layout

As mentioned previously, the primary objective of the reporting tool is to provide direct visualization and edition of raw data. Since this tool does also play an important role in terms of navigating and communicating information in a more textual way (as compared to the pictograms, diagrams and maps that drive the systemic and spatial logics), raw data need to be used as the basis upon which to build written and illustrated reports. The rich description framework provided allows added value to be given to the data themselves.

After an overall description of the reporting tool’s interface, focus is given to some of its most remarkable features, i.e. dynamic update and filtering, heritage and polymorphism, and navigation facilities. Some final comments and perspectives form the conclusion.

# General layout: catalogue and report sheets

As shown in figure 2, the user interface is split in two areas: a left banner hosting some base tools and a main view area hosting a collection of elements or items in a tab control.

The left banner provides an element library which exposes the whole set of system elements forming the ontology used as a basis for the description of the domain addressed; or instance, in the case of urban water management, the ontology defined by Schenk and others (2009). This library reflects the intrinsic systemic nature of the database structure, where any element is schematized either as node, information or interaction (Roquier and others 2010). The applied urban water management reference ontology includes a large variety of nodes (including non structural elements, such as studies, strategies, legal references, policies, funds or stakeholders, and structural elements, among which (I) natural elements, such as rivers, ecosystems, or watersheds, (II) physical resources, such as surface waters, or groundwater, but also flora and fauna or energy, and (III) anthropogenic elements, such as water and sanitation infrastructures, or on water equipment). This ontology also includes a set of information items (of different types, such as texts, numeric values, Booleans, geometries, and all kind of files, including pdf files, images, audio and video streams, program executables, etc.) and a collection of  interaction types (logical connections – instantiation, group membership, specialization, property, etc. – and material fluxes – energy, money, water, material, data). Although these elements are primarily used by the system viewer (to draw diagrams depicting the systemic layout of various water management related issues for instance), accessing readily their characteristics and related information with the reporting tool is a fundamental feature. This is achieved by dragging an element from the library and dropping it into the main area, causing its report sheet to be retrieved from the database and added as a new tab in the general, uppermost, tab item collection. This drag and drop process applies as well with results from the simple search tool or elements dragged from the system or the geographic viewers.

The main area shows basically a tab collection, each tab item hosting the information set related to an individual element (Birmingham's Frankley water treatment work and Minworth waste water treatment plant in the example depicted by fig. 2). Each item's information set consists actually of a title bar, a collection of report sheets along with a data catalogue.

The title bar hosts the pictogram of the class the selected item belongs to, along with the item's name and a tiny expander button that provides quick access to a small subset of summary characteristics, such as class, ID, status, etc. The tab control located just below hosts the group of report sheets themselves. These report sheets may include several distinct reports related to the same item, for instance a report on the WWTP technical specificities along with another one listing its staff members and their duties, etc.

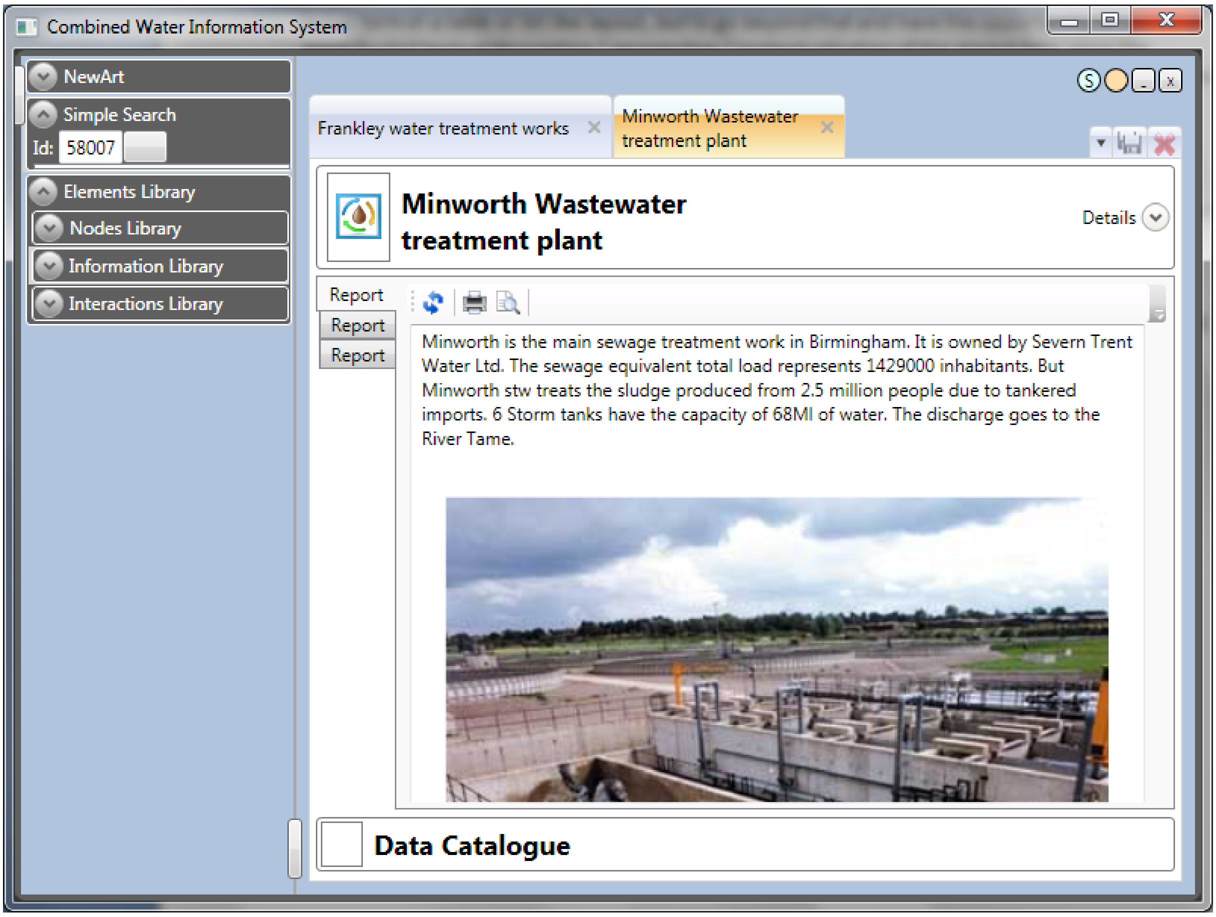



Figure 2: ART user interface

The data catalogue found at the bottom of the area can be lifted up and down, to give access to its content (figure 3). Currently the data catalogue presentation layout provides the complete list of properties (which are themselves system elements) that have been recorded to characterize the current system element or item (as can be seen for the WWTP, the catalogue can be used to access every single data item, as for instance the observed N concentration or BOD values). This list includes properties, interactions, group memberships, and themes for an instance – with additional items for

classes such as taxonomy, i.e. parent and child classes, class properties and interactions, class instances, etc.

Edition of the report sheets is supported by a markup language (figure 3) that can accommodate a wide range of dynamic controls, providing the necessary support to include the various kinds of information supported by the database (texts, values, images, hyperlinks, etc.) in the actual reports.

Raw information taken from the data catalogue can be inserted in a report by a simple drag and drop operation whereas using a markup language allows the fine tuning of a report's layout.

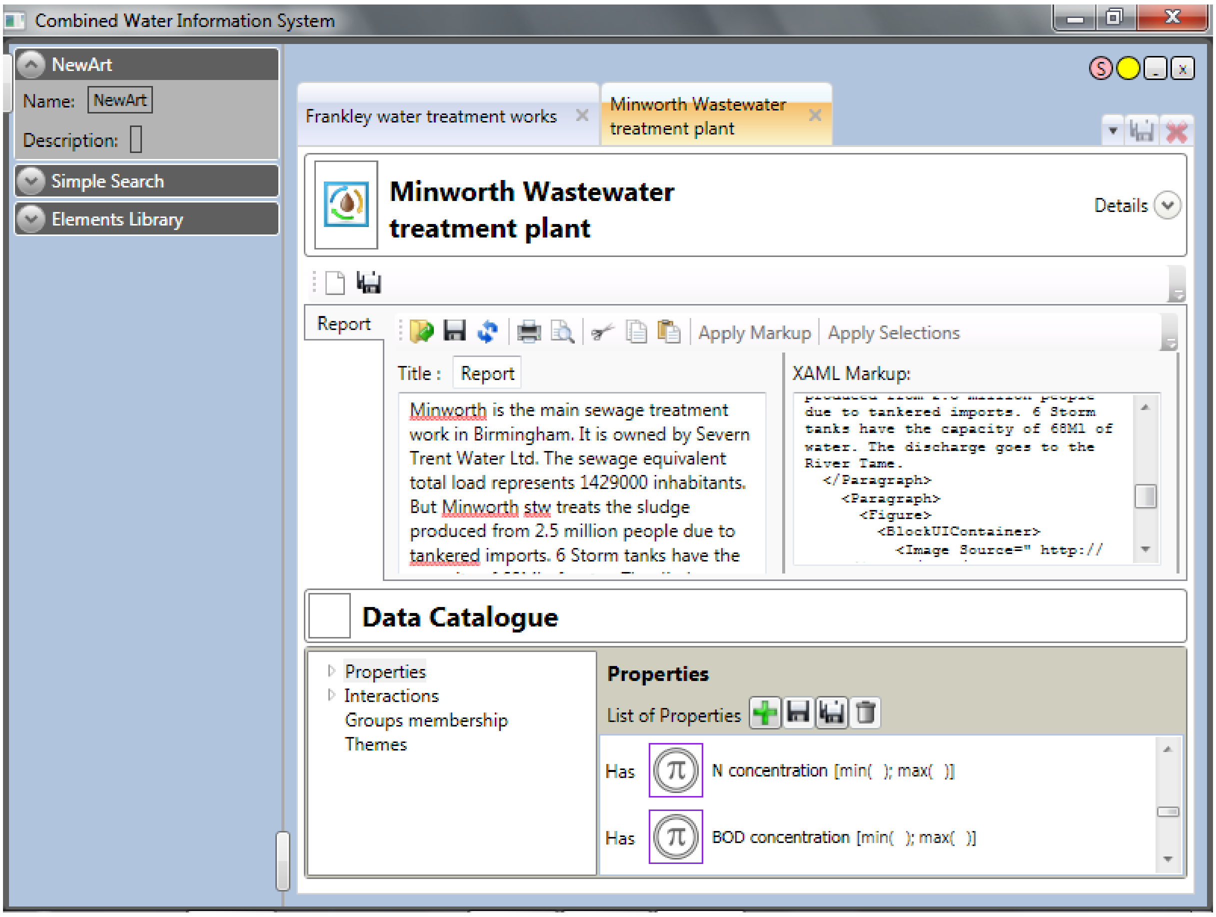



Figure 3: Data catalogue and report edition mode

# Dynamic update and filtering

Through insertion of data catalogue items in the report sheets, linkages between report sheets and the data items they contain are created. As a consequence, any change to the stored data is automatically reflected on the linked report sheets.

In the case for instance of a series of numerical values, such as the BOD of the waste water treatment plant outflow, this dynamic update link means that if a report sheet references the current BOD value, the latest recorded value will actually be shown, ensuring that the report sheet does always provide the most up-to-date information available in the database.

As a matter of fact, a report sheet does not need to reference the latest registered value, but might also be linked to the values recorded at a given moment or within a given time frame. This time basis filtering process allows for recreating the report sheet as produced back in time.

Furthermore, filtering functionalities may also be applied to another important dimension of a dataset, the thematic dimension. Since the GenDSS is meant to provide support for scenario-based planning, the data related to various alternative strategies need to be stored in the database. A thematic is therefore the group of data items related to a given future strategy, reality being just one special case of thematic.

# Inheritance and polymorphism

As mentioned previously, the database is structured on a systems basis, according to the system modeling language SYSMOD (Roquier and others 2010). This language defines system elements along two axes: the type and the level. The type can be either “node”, “interaction” or “information”. The level can be either “class”, “instance” or “property”. A “node” is a system element, such as a stakeholder, or a river. An “interaction” is a functional relation between two groups, such as a flux or an influence. An “information” is an actual value such as a numeric data or a text.

As a consequence, actual elements are instances of a class. Classes may inherit some or all features from one or several other classes. Moreover two classes sharing the same parent class may have the same inherited method or characteristic, but with appropriately different results or outputs.

These features known as inheritance and polymorphism have an interesting practical consequence regarding dynamic reporting: they allow a system element's report view to contain not only its own report sheets, but also the report sheets of its parent classes. For instance, in the active reporting view, Birmingham's Minworth Waste Water Treatment plant comes up with a self-descriptive report sheet. Additionally, the more generic report sheet summarizing waste water treatment techniques linked to the waste water treatment plant class is also provided, along with the report sheet describing the waste water treatment legal framework, a given WWTP also inherits from.

# Data navigation

The ART tool provides means to access raw data as well as more elaborated reports. Its role is however also to complement the systemic and geographic viewers in offering an alternative way to browse and navigate information.

This data navigation is provided both at the data catalogue level and at the report sheet level. At the data catalogue level, navigation is achieved simply by clicking the system elements referenced in the catalogue list, operation that will retrieve the selected items' own report sheets and data catalogue and add them to the ART instance's tab collection in a new tab item.

The report sheets themselves offer a more direct way of browsing information in that they can host any kind of catalogue items (system elements) in an hyperlink-like structure. Again a simple click on one of these "hyperlinks" will retrieve and show the selected items own report sheets, thus offering the opportunity to "jump" from one report sheet to another in Web-browser like way. Obviously real Web links opening a Web browser to the referenced Web page are also functional.

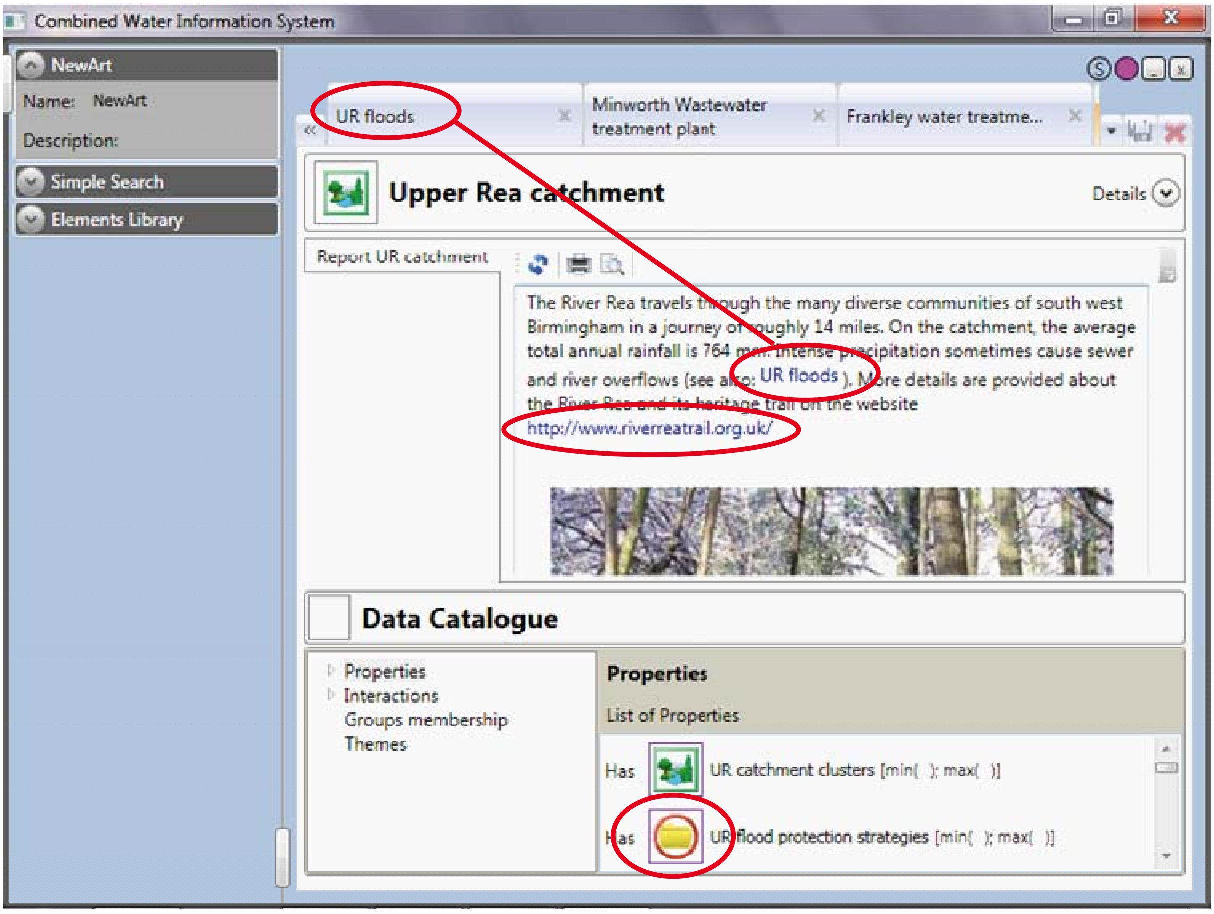



Figure 4: Examples of data navigation possibilities

# Conclusions and perspectives

The examples presented in this paper demonstrate that ART is a powerful dynamic reporting tool, offering a rich environment to elaborate and communicate reports that allow one to add value to data stored in a database.

Taking advantage of the systemic nature of the database it is built upon, ART offers a set of uncommon original additional features (especially inheritance and polymorphism, and navigation facilities) that turn it into an essential piece of the combined information system it belongs to.

Along with its complementary systemic and geographic navigation tools, ART contributes to provide the versatile and flexible environment to be used as an information and knowledge sharing platform and support scenario based planning.

The current data catalogue layout, although very handy for developers or end-users familiar with system logics, might be a bit uncomfortable for a non-specialized end-user or with respect to data edition. Therefore, other presentation layouts, for instance more classical tabular presentations to edit properties of grouped items, will be developed in the near future. They might put more emphasis on intrinsic properties and less on interaction links, and eventually form the core of an additional autonomous CWIS module.

# References

Eclipse Foundation (2009) Eclipse BIRT home. Accessed online 02/01/2010: http://www.eclipse.org/birt/phoenix/.

GrapeCity (1997) ActiveReports Reporting and Business Intelligence (BI) tools for Microsoft Visual Studio - GrapeCity. Accessed online 02/01/2010: http://www.datadynamics.com/.

Jaspersoft (2000) Embeddable Java Reporting Library - JasperReports. Accessed online 02/01/2010: http://www.jaspersoft.com/jasperreports.

Kantardzic M (2003) Data Mining: Concepts, Models, Methods, and Algorithms. Wiley, IEEE Press, 360 pp.

Ni Q, Yarlagadda PKDV, Lu WF (2007) A configuration-based flexible reporting method for enterprise information systems. Computers in Industry 58 (5): 416-427.

O'Brien J, Marakas G (2009) Management information systems. McGraw-Hill / Irvin, Boston MA, 624 pp.

Pedersen TB, Jensen CS (2001) Multidimensional database technology. Computer 34 (12): 40-+.

Roquier B, Schenk C, Soutter M, Mermoud A (2010) SYSMOD: A systems modelling language for environmental information. Submitted to Ecological Modelling.

SAP (2009) Crystal Reports Solutions. Accessed online 02/01/2010: http://www.sap.com/solutions/sapbusinessobjects/sme/reporting-dashboarding/reporting/index.epx.

Schenk C, Roquier B, Soutter M, Mermoud A (2009) A system model for water management. Environmental Management 43 (3): 458-469.

Schenk C, Roquier B, Soutter M, Mermoud A (2010a) A generic decision support system. In preparation.

Schenk C, Roquier B, Soutter M, Mermoud A (2010b) A systems-based information system. Submitted to Information Systems.

Shadan M (2005) Entreprise dashboards: design and best practices for IT. John Wiley & Sons, New York, 222 pp.

# ANNEX III

## A SYSTEMS-BASED INTEGRATED MODELLING FRAMEWORK

# Preface

This annex presents how models can be linked to the GenDSS, in order to export data to run simulations and import back the results.

This annex is adapted from a paper in preparation, written by Roquier B, Schenk C, Soutter M, Mermoud A (2010) and which will be entitled: A systems-based integrated modelling framework.

# Abstract

To support the management of environmental resources, environmental integrated modelling frameworks (EIMFs) aim at promoting the reuse of data and models in order to save time and money during the creation of decision support systems (DSSs), while improving their quality.

As an alternative to programming frameworks, this paper presents a new systems-based integrated modelling framework (SIMF) designed for a specific software environment: the GenDSS, a generic tool to manage and consult systems' data, and to link models. SIMF allows users to adapt the GenDSS software to specific tasks by integrating ontologies of knowledge domains and by configuring features of the application. Therefore, using SIMF with the GenDSS reduces the needs for software development compared to other DSSs built with conventional EIMFs.

# Introduction

Environmental integrated modelling frameworks (EIMF) have been developed as a support to the management of environmental resources. They aim at promoting the reuse of data and models in order to save time and money during the creation of decision support systems (DSSs) or integrated assessment tools (IATs), while improving their quality (Rizzoli and others 2008).

In the environmental field, most recent integrated modelling frameworks (IMFs) consist of suites of software engineering tools dedicated to the development of new DSSs or IATs. They propose modelling frameworks that are generally combined with some application programming interfaces (API). Examples of such EIMFs are The Invisible Modelling Environment (TIME) (Rahman and others 2003, Rahman and others 2004), Tornado (Claeys and others 2006), the Object Modelling System (OMS) (David and others 2002), JAMS (Kralisch and Krause. 2006), ModCom (Hillyer and others 2003) or OpenMI (Gregersen and others 2007). Other IMFs, such as MATLAB (The MathWorks Inc. 2009) or Mathematica (Wolfram Research Inc. 2009) frameworks, are numerical computing environments using dedicated programming languages and providing rich libraries of components. They facilitate the development of DSSs (among many other things) based on predefined objects such as matrices, algorithms or charts. A third group of IMFs comprises specific software environments that supports models integration, such as the Modular Modelling System (MMS) (Leavesley and others 1996), the Dynamic Integration Architecture System (DIAS) (Sydelko and others 2001), the Interactive Component Modelling System (ICMS) (Rizzoli and others 1998, Reed and others 1999), Tarsier (Watson and Rahman 2004) or the Spatial Modelling Environment (SME) and its associated module specifications (Maxwell 1999, Voinov and others 1999, Voinov and others 2004).

As stated by many authors (Ceccaroni and others 2004, Rizzoli and others 2005, Scholten and others 2007, Villa 2007), the use of ontologies is an effective way to support the integration of models in systems modelling, with some benefits as regards the rigor and consistency of the process. In this perspective, Rizzoli et al. (2008) propose a generic architecture for EIMFs which makes use of ontologies along with models, workflows and other tools. This generic architecture is dedicated to EIMFs as programming frameworks and therefore cannot be applied as such to software environments.

As an alternative to programming frameworks, this paper presents a new systems-based IMF architecture for software environments, which is an adaptation of the generic architecture for EIMFs (Rizzoli and others 2008) for a specific software: the GenDSS (Schenk and others 2010b), a generic tool to manage and consult systems' data, as well as to link models to run simulations. This system-based integrated modelling framework (SIMF), like the GenDSS, is build on the basis of a particular modelling language – SysMod (Roquier and others 2010) – which allows modellers to create ontologies as well as to model systems. Taking advantage of the GenDSS modules, a new approach for models integration is illustrated in the paper.

# Framework architecture

Adapted from the generic architecture for environmental integrated modelling frameworks (EIMFs) (Figure 1.a) proposed by Rizzoli et al. (2008), the systems-based integrated modelling framework (SIMF) has been formalized to support the coupling and integration of models within the GenDSS environment (Schenk and others 2010b). Contrary to the definition of EIMF proposed by Rizzoli, which considers EIMF as a suite of software engineering tools for the development of DSSs or IATs, SIMF isn't a programming framework for the development of new softwares. SIMF is a conceptual framework implemented into the GenDSS application to further enable models integration. For that reason, the architecture of SIMF differs from the generic architecture for EIMFs, as it is based on a unique, generic software which can be adapted to specific tasks through the use of ontologies, models, configuration tools and other features (Figure1.b).

The GenDSS (Schenk and others 2010b) is a generic tool that can be applied to any kind of systems, be they environmental, economic, social or political. Its genericity comes from the adoption of a particular modelling language – the SysMod language (Roquier and others 2010) – which allows both the creation of ontologies and the modelling of systems in a common frame. The GenDSS requires also a specific database, whose core structure reflects the constructs of the SysMod language and thus doesn't need to integrate heterogeneous databases and data structures (Schenk and others 2010a).

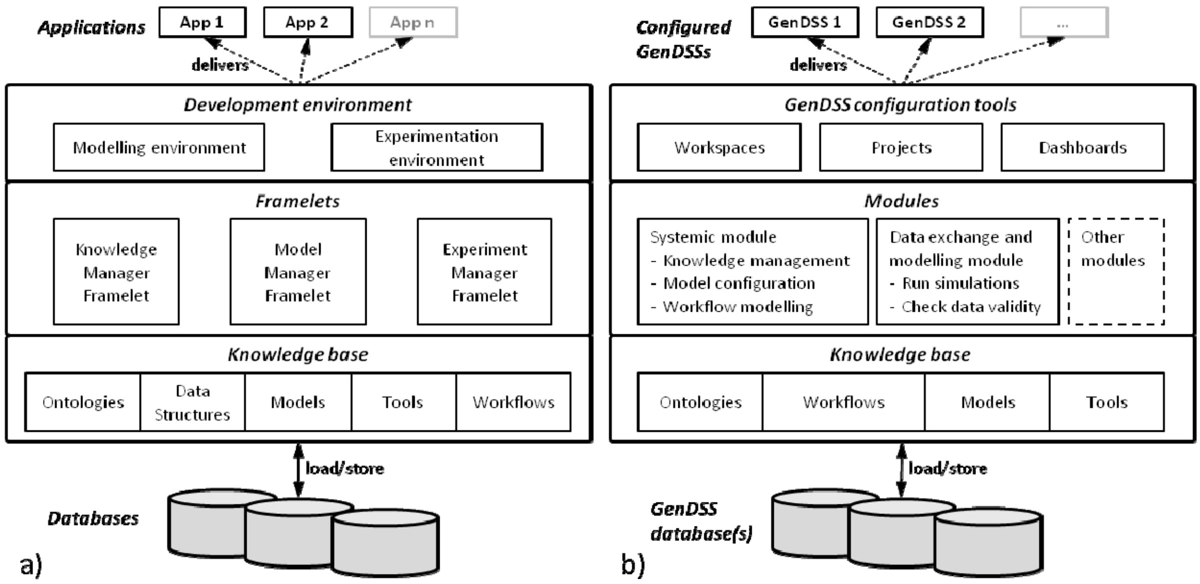



Figure 1: a) The generic architecture for environmental integrated modelling frameworks (Rizzoli and others 2008) and b) the SIMF architecture

# The knowledge base

The knowledge base is the fundamental layer of the SIMF architecture and is populated by accessing the content of the GenDSS database(s). It contains ontologies, workflows, models and tools which are necessary components to operate the GenDSS.

### Ontologies

In computer science, an ontology is a formal representation of a knowledge domain. It consists of a set of organized concepts, properties and relationships. There are many ontology languages available, such as the Web Ontology Languages (OWL), DAML+OIL or the Knowledge Interchange Format (KIF). In this paper, the SIMF ontologies are based on the SysMod language; an ontology language specifically designed for systems modelling (Roquier and others 2010).

As stated by Rizzoli et al. (2005), "the use of ontologies is advantageous as it (a) supports the automatic generation of code templates for models and domain classes in different Integrated Modelling Frameworks, (b) it facilitates the application of a reasoner (inference engine) on the structured knowledge, which can detect abnormalities or conflicts in model interfaces, and (c) it supports model linking in a content enriched way, which can be proven valuable for avoiding common problems related to poor semantics of model interfaces."

As the GenDSS database structure reflects the constructs of the SysMod language, the latter is the only ontology language that is compatible with the GenDSS environment. There is however the possibility to convert ontologies from other formats in SysMod.

### Workflows

In the context of integrated modelling frameworks, workflows describe series of activities and computations that enable complex modelling and data analysis (Ludascher and others 2004). The integration of different models is made, for instance, through their representations in a scientific workflow which can be then parsed and run to perform simulations.

Compared to other frameworks, SIMF has the particularity of using the same modelling language (SysMod) to model ontologies, workflows, and any other systems. The result is a saving in terms of IT development, given that these different modelling activities can be performed in a single environment.

### Models

The GenDSS database can store models. To be used in the system, these models need an equivalent class in the ontology. This means the ontology contains a generic class "model" that has several subclasses representing the specific models and their properties. These model classes can be instantiated into some workflows.

### Tools

Beside models, the Knowledge base contains tools that are used to manage and transform data. Tools are required, for instance, to convert values from one unit to another unit or to convert file from a format to another format. In the context of the GenDSS, the integration of models needs such tools to convert the objects selected as inputs into the format required by the model and conversely, to transform the output file in a set of “system elements” expressed in the SysMod language.

# Uses of the GenDSS modules

As stated above, SIMF isn’t a programming framework designed for the creation of new software applications; it is rather a modelling framework that uses the GenDSS as a modelling environment. Therefore, compared to the generic architecture proposed by Rizolli et al. (2008), SIMF doesn’t have framelets (some lightweight and highly specialised frameworks), but uses instead the different modules of the GenDSS to access the Knowledge base. Among the modules of the software, two of them are relevant to integrated modelling: the Systemic Viewer and the Data Exchange and Modelling Module. The Systemic Viewer allows the GenDSS users and modellers to edit systems’ data, configure the inputs and outputs of models, and create workflows integrating data, tools and models. On the other hand, the data exchange and modelling module provides support to import/export data, as well as to perform simulations with some models externally linked to the GenDSS.

### Knowledge management

Data edition is performed using the GenDSS modules. While the Geographic Viewer is designed for the edition of spatial attributes and the Active Reporting Tool for editing the other information and creating interactive pages that display values, images or hyperlinks from the database, the Systemic Viewer is a cornerstone of SIMF as it is needed to edit ontologies, represent knowledge from a systemic perspective, define model configurations and create scientific workflows.

Basic knowledge management is illustrated by the Figure 2, where environmental information is handled using the Systemic Viewer. The figure depicts a study area consisting of subcatchments, themselves grouping geographic clusters. It represents a neighbourhood of the city of Birmingham, which is used as a case study for running hydrological simulations with the help of the GenDSS.

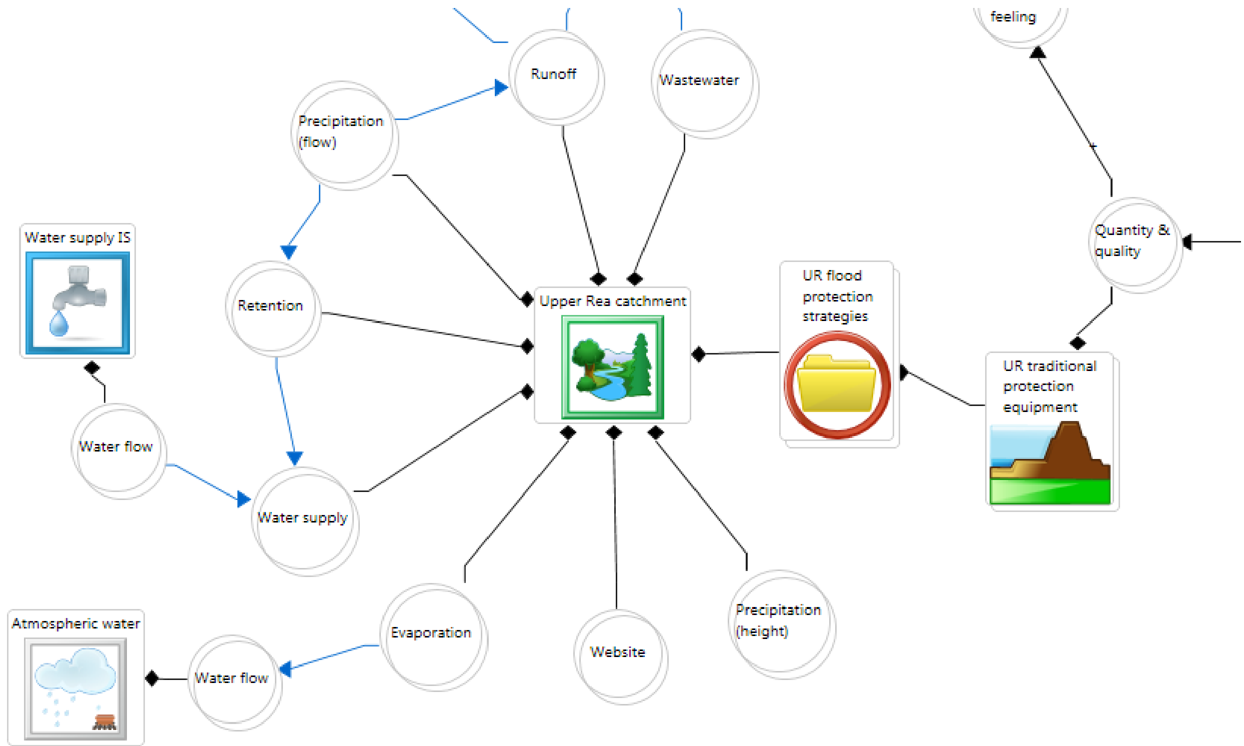



Figure 2: Environmental information expressed in diagrams with the Systemic Viewer

## Model configuration

The Systemic Viewer also allows modellers (or model designers) to define model inputs and outputs through the help of "Model Metaviews". Figure 3 shows, for instance, a subset of the inputs required by a hydrological model. The underlined elements are classes; they are part of the ontology and can be utilized to check the validity of models inputs/outputs. The other elements, displayed in the boxes with shadows, are groups of objects or information that are used as "container" for selecting and storing the inputs/outputs of the model.

The elements listed in these Model Metaviews are used by the Data Exchange and Modelling Module to create lists of the models inputs/outputs and check the validity of the data selected by the user by

confronting them with the expected class (defined in the Metaview). This Data Exchange and Modelling Module works as a shopping basket, providing a list of inputs/outputs needs based on the content of the "Model Metaviews" (Figure 4).

As regards the generic architecture proposed by Rizzoli et al. (2008) in Figure 1, it is worth highlighting that the System Viewer and the Data Exchange and Modelling Module fulfil the same functions as the Model Manager Framelet.

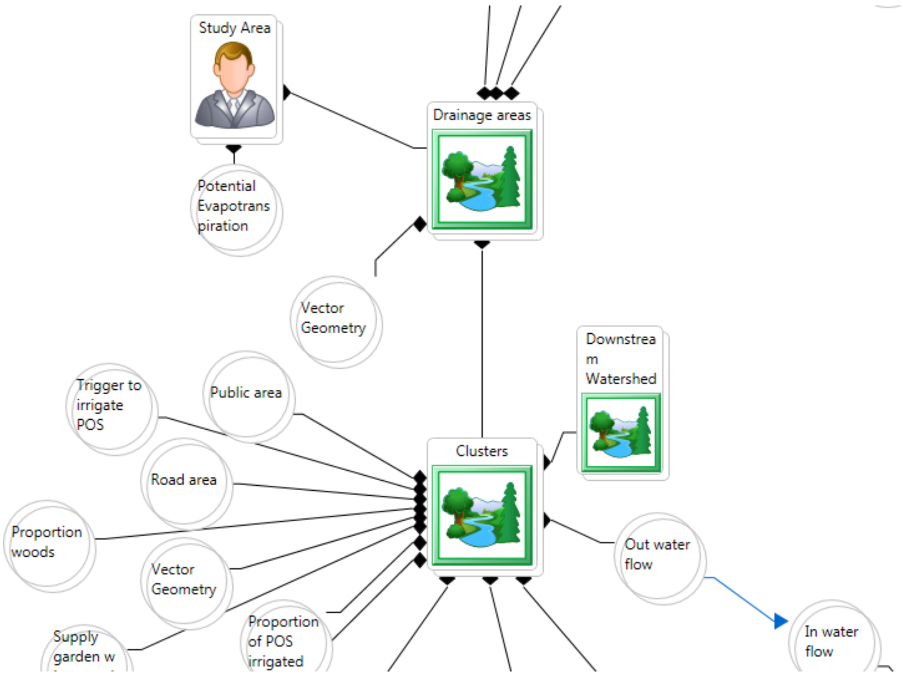



Figure 3: An example of a Model Metaview

## Modelling workflows

The Systemic Viewer displays information according to the systemic structure imposed by the SysMod language. At a basic level, SysMod allows modelling systems that describe elements and processes of the real world (as shown in Figure 2) or, at a meta-level, it also enables creating workflows encompassing data sources, tools and models. These resulting workflows are to be using models and tools available from the Knowledge base to perform activities such as simulations involving several coupled models. The example of the Figure 5 shows the integration of the model mentioned in the previous section with tools to convert the "SysMod" objects selected in the GenDSS into the input format of the model and then convert back the outputs into the SysMod format. The second part of the workflow illustrates the possible use of other kind of models such as basic

mathematical functions like the "Annual average", calculating for instance an average based on series of daily data.

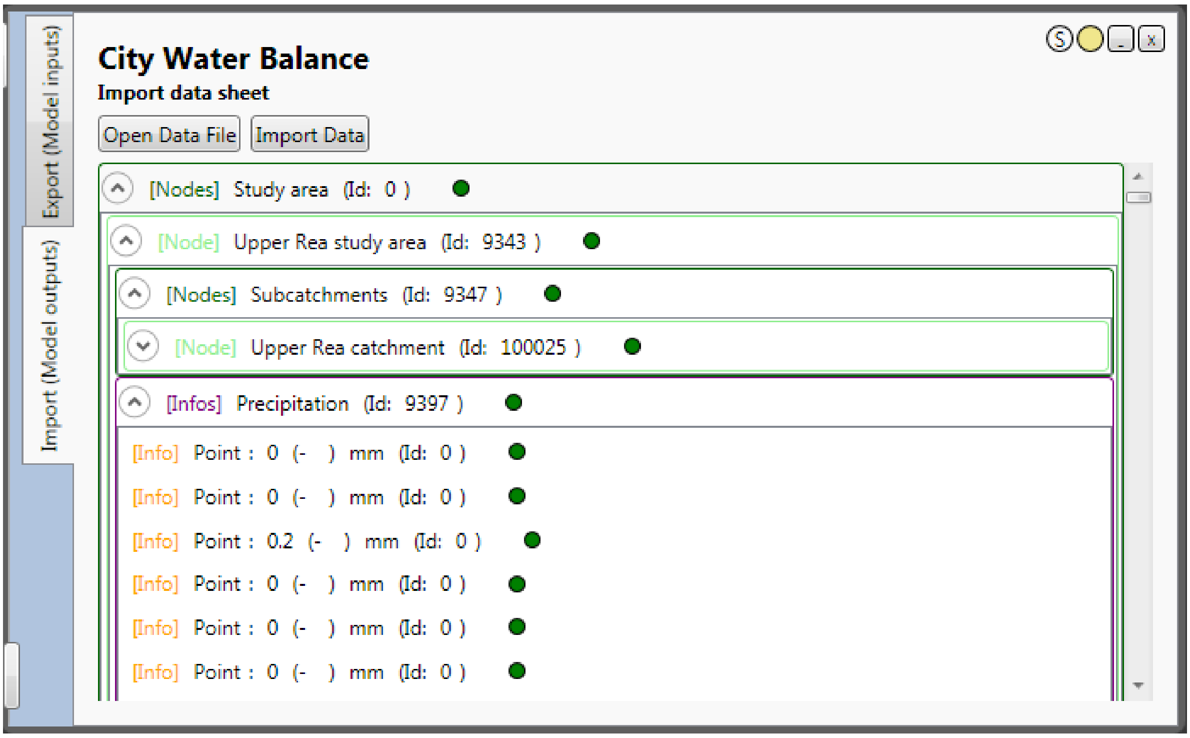



Figure 4: The Data Exchange and Modelling Module

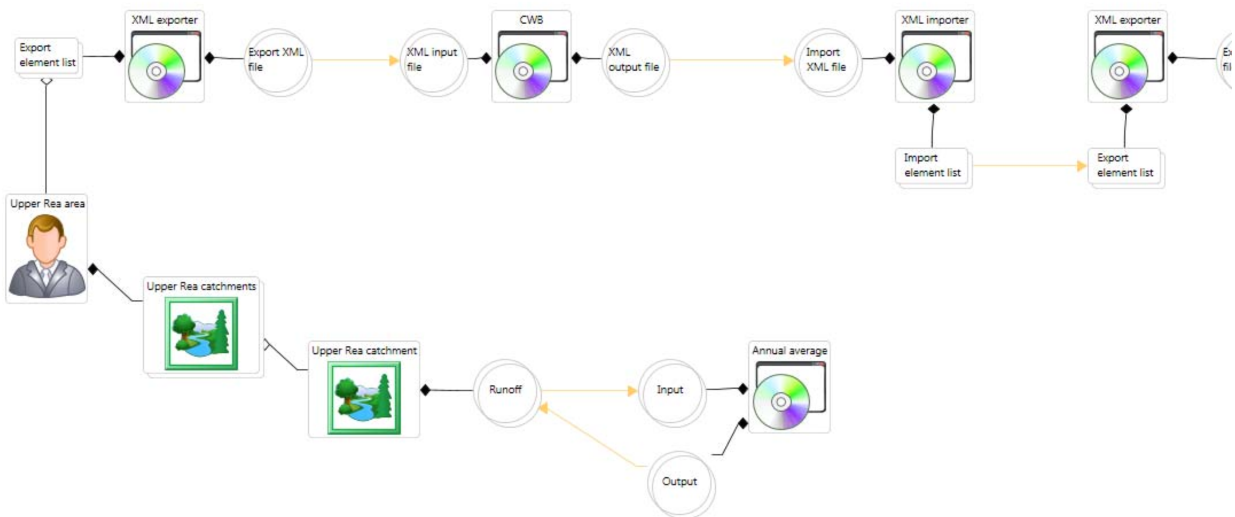



Figure 5: Workflow showing the integration of a hydrological model (CWB) coupled with XML translation tools

In addition to the activities of knowledge management and model configuration, the Systemic Viewer is, therefore, also the tool to be used for models integration. This function corresponds to the third framelet of the generic architecture proposed by Rizzoli et al. (2008): the Experiment Manager Framelet.

# Configuration tools

Unlike the Modelling and Processing environments of the generic architecture for EIMFs -which are software environments that facilitate the development of applications (Figure 1) - the SIMF cannot be used to create new softwares. The SIMF is indeed only a framework defined within the GenDSS software. However, the GenDSS can be adapted and configured at different levels to match the needs of the users. Firstly, at a knowledge management level, implementing an ontology of a knowledge domain allows the GenDSS to use a vocabulary that suits the needs of the users. Then, from a programming perspective, the GenDSS is based on a modular architecture (Schenk and others 2010b) and therefore could be extended with other modules providing extra functionalities. Finally, the GenDSS offers several tools to configure the GenDSS: (1) The workspaces used to define the layout of the applications by sharing the main window in several regions and assigning them different modules/views. (2) The projects that bring together some sets of workspaces and views, in order to facilitate the access to preselected information. (3) The dashboard, which is a “welcome” panel configured by the GenDSS administrator, that allows (new) users to directly access some projects, workspaces or views without being forced to go through the steps of database configuration and connexion.

# Integrated modelling with the GenDSS

Model integration is possible in different ways depending on the form of the model, the requirement for model specific user inputs that are external to the GenDSS and the form of the input/output data structures applicable to the model. The next section describes briefly the general principles for exporting, importing and interacting with models within SIMF (Figure 6). The steps are:

1. From the user interface, the user selects the data to be exported. These data are summarized in the Data Exchange and Modelling view. To guide the user during this process, the view contains a definition of the expected inputs. The latter are clearly identified in the "model meta-views" using a schematic representation of the model inputs.
2. Upon completion of the selection process, the data can then be exported as a SysMod-format xml file.
3. This file can be transformed (through an intermediate data translation routine) into specific formats, such as CSV, Excel or proprietary model input files, if required. Before this stage, the file is first checked and possible problems are notified back to the Data Exchange and Modelling view. If the file is adequate, only then will it be transformed into the target format. (Step 2-3 can be performed with a single tool).
4. The target (translated) files are used as input for the model.
5. The model is then run and model output files are created.
6. The model output file(s) is then transformed back into a SysMod-format xml file (SysMod-import file).
7. This SysMod import file can be read by the GenDSS and shown in the Data Exchange and Modelling view.
8. Finally, the results may be displayed back in the GenDSS various views and stored in the database.

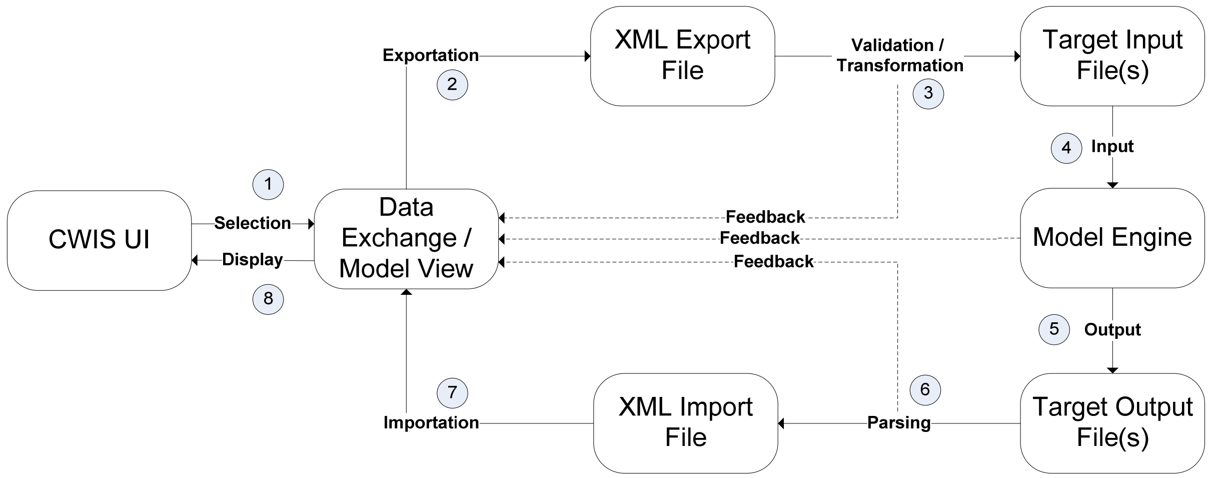



Figure 6: Overview of the data exchange and model interoperability process within SIMF

# Conclusion

The systems-based integrated modelling framework (SIMF) isn't a software engineering environment and therefore cannot be used to create new applications with customized visual interface. However, implemented into the GenDSS environment, SIMF allows users to adapt the GenDSS software to a specific task by integrating an ontology of the adequate domain and by configuring and using specific software features such as workspaces to define the layout of the application.

While conventional EIMFs allow programmers to create personalized DSSs that perfectly match users' expectations in terms of content and visual aspects, SIMF (implemented into the GenDSS) offers a generic modelling environment that handle any kind of content, but with an imposed visual interface. Therefore, because of its genericity, the GenDSS may be less adequate to deal with specific issues than other specialized DSSs. On the other hand, the GenDSS provides a general and evolutionary framework, which allows avoiding the creation and integration of a multitude of disparate tools.

# References

Ceccaroni L, Cortés U, Sànchez-Marrè M (2004) OntoWEDSS: augmenting environmental decision-support systems with ontologies. Environmental Modelling & Software 19 (9): 785-797.

Claeys F, Pauw DJWD, Benedetti L, Nopens I, Vanrolleghem PA (2006) Tornado: A versatile and efficient modelling & virtual experimentation kernel for water quality systems. In: Voinov A, Jakeman AJ and Rizzoli AE (ed) Proceeding of iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software", Burlington, USA.

David O, Markstrom SL, Rojas KW, Ahuja LR, Schneider IW (2002) The object modeling system. In: (ed) Agricultural System Models in Field Research and Technology Transfer, Boca Raton, pp 317-330.

Gregersen JB, Gijsbers PJA, Westen SJP (2007) OpenMI: Open modelling interface. Journal of Hydroinformatics 9 (3): 175-191.

Hillyer C, Bolte J, van Evert F, Lamaker A (2003) The ModCom modular simulation system. European Journal of Agronomy 18 (3-4): 333-343.

Kralisch S, Krause. P (2006) JAMS – A Framework for Natural Resource Model Development and Application. In: Voinov A, Jakeman AJ and Rizzoli AE (ed) Proceeding of iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software", Burlington, USA.

Leavesley GH, Markstrom SL, Brewer MS, Viger RJ (1996) The modular modeling system (MMS) — The physical process modeling component of a database-centered decision support system for water and power management. Water, Air, & Soil Pollution 90 (1): 303-311.

Ludascher B, Altintas I, Berkley C, Higgins D, Jaeger E, Jones M, Lee EA, Tao J, Zhao Y (2004) Scientific workflow management and the Kepler system. In: (ed) Proceeding of GGF Workshop on Workflow in Grid Systems, Berlin, GERMANY, pp 1039-1065.

Maxwell T (1999) A paris-model approach to modular simulation. Environmental Modelling and Software 14 (6): 511-517.

Rahman JM, Seaton SP, Cuddy SM (2004) Making frameworks more useable: using model introspection and metadata to develop model processing tools. Environmental Modelling & Software 19 (3): 275-284.

Rahman JM, Seaton SP, Perraud JM, Hotham H, Verrelli DI, Coleman JR, Mssanzi (2003) It's TIME for a new environmental modelling framework. In: (ed) Proceeding of International Congress on Modelling and Simulation, Townsville, AUSTRALIA, pp 1727-1732.

Reed M, Cuddy SM, Rizzoli AE (1999) A framework for modelling multiple resource management issues--an open modelling approach. Environmental Modelling and Software 14 (6): 503-509.

Rizzoli AE, Davis JR, Abel DJ (1998) Model and data integration and re-use in environmental decision support systems. Decision Support Systems 24 (2): 127-144.

Rizzoli AE, Donatelli M, Athanasiadis IN, Villa F, Huber D (2005) Semantic links in integrated modelling frameworks. In: (ed) Proceeding of International Congress on Modelling and Simulation (MODSIM05), Melbourne, AUSTRALIA, pp 412-423.

Rizzoli AE, Leavesley G, Ascough JC, Argent RM, Athanasiadis IN, Brilhante V, Claeys FHA, David O, Donatelli M, Gijsbers P, Havlik D, Kassahun A, Krause P, Quinn NWT, Scholten H, Sojda RS, Villa F (2008) Integrated modelling frameworks for environmental assessment and decision support. In: Jakeman AJ, Voinov AA, Rizzoli AE and Chen SH (ed) Environmental Modelling, Software and Decision Support, pp 101-118.

Roquier B, Schenk C, Soutter M, Mermoud A (2010) SYSMOD: A systems modelling language for environmental information. Ecological Modelling Submitted.

Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A (2010a) A systems-based, generic environmental data model. Environmental Management Submitted.

Schenk C, Roquier B, Soutter M, Mermoud A (2010b) A systems-based generic decision support system. In preparation.

Scholten H, Kassahun A, Refsgaard JC, Kargas T, Gavardinas C, Beulens AJM (2007) A methodology to support multidisciplinary model-based water management. Environmental Modelling & Software 22 (5): 743-759.

Sydelko PJ, Hlohowskyj I, Majerus K, Christiansen J, Dolph J (2001) An object-oriented framework for dynamic ecosystem modeling: application for integrated risk assessment. Science of the Total Environment 274 (1-3): 271-281.

The MathWorks Inc. (2009) MATLAB - The Language Of Technical Computing. Accessed online 11/17/2009: http://www.mathworks.com/products/matlab/.

Villa F (2007) A semantic framework and software design to enable the transparent integration, reorganization and discovery of natural systems knowledge. Journal of Intelligent Information Systems 29 (1): 79-96.

Voinov A, Costanza R, Wainger L, Boumans R, Villa F, Maxwell T, Voinov H (1999) Patuxent landscape model: integrated ecological economic modeling of a watershed. Environmental Modelling and Software 14 (5): 473-491.

Voinov A, Fitz C, Boumans R, Costanza R (2004) Modular ecosystem modeling. Environmental Modelling & Software 19 (3): 285-304.

Watson FGR, Rahman JM (2004) Tarsier: a practical software framework for model development, testing and deployment. Environmental Modelling & Software 19 (3): 245-260.

Wolfram Research Inc. (2009) Mathematica. Accessed online 11/17/2009: http://www.wolfram.com/products/mathematica/.

# CURRICULUM VITAE

# Colin SCHENK

Swiss citizen
Born on the 10th of December 1976

## Education

| | |
|---|---|
| 2010 | PhD in Hydroinformatics<br>Ecole Polytechnique Fédérale de Lausanne (EPFL)<br>Ecohydrology Laboratory |
| 1999 | Diplôme d'ingénieur en Sciences et Ingénierie de l'Environnement (Master degree in Environmental Sciences and Engineering) |
| 1994 | Maturité fédérale scientifique (type C)<br>Hoffmann-La Roche award for the chemistry research work<br>Gymnase cantonal de La Chaux-de-Fonds |

## Professional experience

| | |
|---|---|
| 2003-2010 | Scientist and PhD candidate (PhD started in 2006)<br>Ecole Polytechnique Fédérale de Lausanne (EPFL)<br>Hydram and Ecohydrology Laboratory (laboratory renamed in 2008) |
| 2002-2003 | Engineer and industrial programmer<br>ASER automation SA |
| 2002 | Founder and manager of a start-up in web technologies<br>Outline SARL, Romania |
| 2001 | Manager of various development projects<br>Centre Culturel Français de Timisoara |
| 2000 | GIS developer for an inter-localities consortium<br>Tura municipality, Hungary |
| 1999-2000 | Manager for starting-up wastewater treatment facilities<br>Degrémont SA (SUEZ environment) |

## Languages

| | |
|---|---|
| French | Mother tongue |
| English | Fluent |
| Hungarian | Fluent |
| Romanian | Intermediate |
| German | Intermediate |

**Computer skills**

| | |
|---|---|
| .NET, C# | Experienced object-oriented application architect and developer |
| Databases | Experienced developer and user |
| Web techs | Fair developer |
| Linux | Fair knowledge |
| C++, Delphi | Fair knowledge |

**Publications**

| | |
|---|---|
| 2010 | Schenk C, Roquier B, Soutter M, Mermoud A. A generic decision support system. In preparation. |
| 2010 | Roquier B, Schenk C, Soutter M, Mermoud A. A systems-based modelling framework. In preparation. |
| 2010 | Soutter M, Roquier B, Schenk C, Mermoud A. ART: An active systems-based navigation and reporting tool. Submitted to Information Systems Frontiers. |
| 2010 | Schenk C, Roquier B, Soutter M, Mermoud A. A systems-based information system. Submitted to Information Systems. |
| 2010 | Schenk C, Roquier B, Brandenberg P, Soutter M, Mermoud A. A systems-based, generic environmental data model. Submitted to Environmental Management. |
| 2010 | Roquier B, Schenk C, Soutter M, Mermoud A. SYSMOD: A systems modelling language for environmental information. Submitted to Ecological Modelling. |
| 2009 | Schenk C, Roquier B, Soutter M, Mermoud A. A system model for water management. Environmental Management 43 (3): 458-469. |
| 2009 | Soutter M, Alexandrescu M, Schenk C, Drobot R. Adapting a geographical information system-based water resource management to the needs of the Romanian water authorities. Environmental Science and Pollution Research 16 (S1): 33-41. |

**Conferences**

| | |
|---|---|
| 2008 | Roquier B, Brandenberg P, Schenk C, Soutter M. City Water: an information sharing platform to support LAs in exploring new strategies. 3rd SWITCH Scientific Meeting, Belo Horizonte, Brazil. |
| 2007 | Schenk C, Soutter M, Roquier B, Mermoud, A. Towards an information system on the water system. 2nd SWITCH Scientific Meeting, Tel-Aviv, Israel. |