An ontological approach to multi-scale modelling of environmental fate and ecological effects in aquatic ecosystems

C. G. Siontorou & F. A. Batzias
Department of Industrial Management & Technology,
University of Piraeus, Greece

Abstract

Aquatic ecosystems are complex and multifaceted webs that continuously shift states to accommodate natural and anthropogenic inputs. Managing such a dynamic scheme necessitates multidisciplinary and multi-scale modelling. Although a variety of modelling strategies and tools have been suggested and used, their applicability is often restricted to one scale (usually the macro- or meso-scale) or to related domains of knowledge (usually hydrogeomorphology or ecology), making accurate predictions on whole system responses difficult to achieve. A critical impediment in constructing an integrative framework lies in the structure of knowledge, which remains context-specific across different disciplines. Acquiring and evaluating information (at the required amount and level of granularity) in collaborative interactions is often a challenge, as is the transfer of this information to others, including scientists, stakeholders, resource managers, policymakers, and the public. This paper presents a methodological framework for bridging the gaps between macro- and micro-scales in ecosystem modelling through an ontological platform designed/developed to accommodate partonomic functions in and between different knowledge domains and levels.
Keywords: aquatic systems, modelling, ontology, knowledge processing, fate of pollutants, denitrification, oil weathering.

1 Introduction

Aquatic ecosystems are among the most complex due to their highly nonlinearity, randomness, and the great variety of interactive multi-processes occurring at multi-scales. Whenever the need arises, e.g., for applying counter-

measures successfully, for performing risk analysis efficiently or for predicting fate and transport of pollutants reliably, modelling becomes indispensable. The highly limited understanding and the very limited measurement data available (or feasible) restrict modelling either to one scale (usually the macro- or meso-scale where measurements are possible [1]) or to related domains of knowledge (usually ecology or hydrogeomorphology where knowledge is available [2]), making accurate predictions on whole system responses difficult to achieve.

Models are, necessarily, simplified representations of the phenomena being studied; a key aspect of the modelling process is the cautious selection of model parameters and assumptions. The optimal model will provide the greatest simplifications for an adequately accurate representation of the processes affecting the system of interest. Inevitably, the optimal model should capture and handle the multi-scale aspect of the system: the substantive, the structural and the dynamic. Developing a multi-scale model is, generally, much more challenging than building a single-scale model, because the modeller has to determine what scales to be involved and how (or where) the involved scales should be connected. Integrating modelling [3, 4] comes at least in three parts: (a) combine models for different processes (especially abiotic with biotic ones) at the same space–time scale, (b) combine processes of the same type across space–time scales (e.g. hydrological transport processes from lateral fluxes in upstream catchments to longitudinal buffering in downstream floodplains, or bioaccumulation from small, practically sessile organisms, to large-scale mobile organisms), and (c) combine processes of different types across different scales (e.g. export of metals from upstream mining areas with exposure in mobile organisms in the downstream parts of the river network).
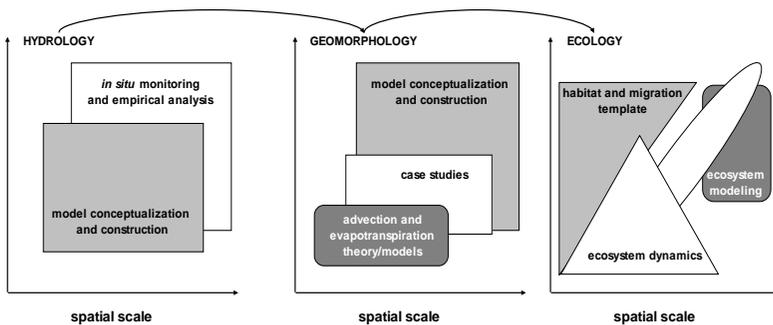
Figure 1: Schematic representation of the knowledge structures of hydrology, geomorphology and ecology; the arrow indicates information granularity level increase.

The participation of experts from a diverse set of scientific disciplines has the potential to serve adequately the first part, since each participant contributes his/her specialized disciplinary knowledge and knowledge acquisition processes that may narrow knowledge gaps. Usually, however, collaborative efforts are loosely organized in a hierarchical way (fig. 1), whereas little attempt is made to

maintain a steady information flow at the required granularity level [5]. The main reason for this lies in the structure of knowledge, which remains context-specific across different disciplines. Acquiring and evaluating information (at the required amount and level of granularity) in collaborative interactions is often a challenge, as is the transfer of this information to others, including scientists, stakeholders, resource managers, policymakers, and the public. The reasons are several and well-known: (i) disciplinary specialization makes it difficult to find a common language in collaborative teams; (ii) existing scientific knowledge may reflect a historical scientific and socio-political context unsuited for the current environmental issues; (iii) mismatches in space and time scales, in forms of knowledge (e.g., macro-systems versus nano-systems), and in levels of precision and accuracy (e.g., qualitative versus quantitative data), may create difficulties in the comparability of results or the flow of information; (iv) disciplinary scopes dictate the way scientists view and study the ecosystems, often favouring certain assumptions over others.

The field of hydrology, for example, is often successful at predicting depositions in relatively small scales where detailed information on impervious area, networks, and topography exists [6]. The geomorphology field is more concerned with wider areas, yet it uses small-scale data for calibrating large-scale models [7]. Ecology, fed with information from hydrological and geomorphological studies, could certainly provide the means to solve environmental problems; yet, it only recently has started to elucidate ecosystem dynamics to produce relatively precise models [1–3], although several knowledge gaps exist, the bridging of which requires the collaboration of many disciplines (such as microbiology and nanotechnology) rendering the field trans-disciplinary [8]. Notwithstanding, the three disciplines show quite different scopes and objectives providing knowledge on different scientific contexts that is not easily collated; moreover, the scale and the levels of precision/accuracy of the hydrology and geomorphology are usually not suited well for integration within an ecological scheme. Consequently, ecological responses to pollution are often considered in qualitative, semi-qualitative or stochastic terms, increasing uncertainty manifold in decision-making [9].

Using an ontological platform can help putting experts' knowledge in the right perspective, providing, also, spatio-temporal projections for the second part of integrating modelling. Ontologies are considered as appropriate modelling structures for representing complex domains under a homogenous terminology and a knowledge integration platform that handles structured or unstructured information resources [10]. Most platforms developed, however, provide a taxonomic (*is-a*) backbone, built on strict hierarchical (often disciplinary) intra-relations, that is used to detect mereological (*part-of*) inter-relations at the same or adjacent knowledge levels (see, e.g., [11–13]), a format that cannot satisfy the third part of integrating modelling. In most aquatic ecosystem ontologies (see e.g., [14–17], the scientific domain (and its scale) is drawn first and then the object of interest is represented as a downstream instance of this domain; there certainly exists the risk either to supersede interrelations existing beyond the

domain's scale boundaries or to rely on partonomic functions that are not consistent with the site hydrogeobiochemistry.

This work presents a methodological framework for bridging the gaps between macro- and micro-scales in aquatic ecosystem modelling through a flexible ontological platform, designed/developed to accommodate partonomic functions in and between different knowledge domains and levels. Focusing on the water system *per se* and its processes between its components, ontology building adopts herein a processualist view of knowledge requirements for capturing the various aspects of this multidisciplinary, dynamic and complex structure: the ontology is built around the sub-system of interest and subsequently integrates the relevant scientific domains (and scales), in order to accurately represent the processes that most likely will take place on site.

2   Methodology

The methodological framework, designed/developed by the authors to support aquatic ecosystem modelling, is briefly described below as a 13-step iterative process.

1. Description of the ecosystem under consideration, putting emphasis on the identification of relevant water-quality issues and concerns.
2. Configuration of an initial (rough) body of multi-/inter-disciplinary knowledge.
3. Selection of experts covering this body of knowledge.
4. Construction of a suitable physical model for the whole system, including compartments and sub-compartments, putting emphasis on the components of each part and their within and between functional interrelation through well defined interfaces (including numerical values of equilibrium constants and kinetic parameters of the transformation functions).
5. Evaluation of the structure of knowledge available and required to serve the needs for information gathering as regards (i) disciplinary forms of available scientific knowledge, (ii) spatial and temporal scales at which that knowledge applies, (iii) precision (i.e., qualitative versus quantitative nature of understanding across different scales), (iv) predictive ecological models, and (v) availability of data to construct, calibrate, and test these models.
6. Specific extraction of information from different knowledge levels of the disciplines involved, putting emphasis on control variables and critical parameters.
7. Reforming of the physical model as an ontological network, where the (biotic and/or abiotic) species are the nodes and the links between them are the equilibrium constants and kinetic parameters, for a static and a dynamic regime, respectively.
8. Mapping of the water-quality issues and their impacts on the ontological network.
9. Informational thinning for eliminating properties/relations of low significance in order to identify the most critical nodes and their inter-relations/interdependencies that relate to the specific water quality problems.

10. Identification and ranking of the strategic nodes, i.e., those having the higher in- and out-degrees.
11. Identification of the strategic control variables and parameters that can be used to verify/quantify/assess the strategic nodes.
12. Construction of a holistic model that can account for all critical processes in all relevant scales.
13. Validation of the model using small-scale measurements of the most important strategic control variables/parameters.

The first step in the recommended framework involves the identification of water-quality issues and concerns associated with the aquatic ecosystem. Regulators, conservation groups, landowners, industries and other stakeholders participate in this process by engaging in focused discussions regarding their interests and needs relative to water resources (a similar procedure has been reported in [18]). It is important to identify key issues and concerns early in the process because such information provides the modeller with a basic understanding of the tradeoffs that need to be addressed by the model.
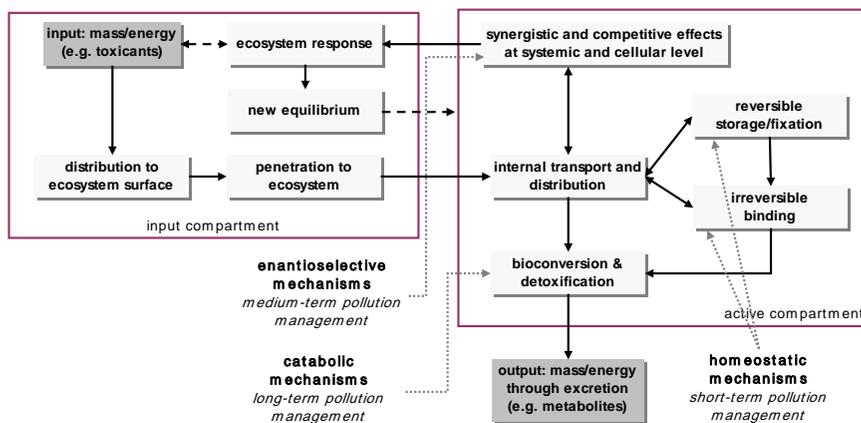
Figure 2: Simplified compartment model for stressor – ecosystem interactions.

In order to perform the partonomic function of step 4, i.e., segmentation of the system into physical compartments and discrimination of each compartment into parts, the ecosystem is considered as an input–output system with boundaries (fig. 2). The inputs to the system are energy and mass; whatever energy or mass enters a defined ecosystem must eventually leave, after some time lag. This is always true, although often the characteristics of the energy and mass leaving the system may differ from those entering, depending upon interactions within the system. Thus, any ecosystem has three primary components to be characterized: the input mass/energy budget, the system itself, and the output mass/energy budget. The output mass/energy from a system can be changed both in character and in timing. Only two mechanisms can change the output: either the input mass/energy budget has changed or the ecosystem

itself has changed. A variety of forces interact dynamically across the time and space scales of the environmental continuum. For any ecosystem domain, this dynamic interplay of forces will be apparent at large variations over time in ecosystem budgets of mass and energy; usually, when budgets are stable there is only one critical variable that greatly dominates the remaining variables /parameters.

Reforming the physical model as an ontological network (step 7) requires the construction of the ecosystem ontology. Since the physical model portrays the ecosystem as a system with inter-related parts (compartments), the ontological network cannot be built on an *is-a* hierarchy as it is based on a partonomy. Therefore, the ontology is built around the (sub)compartment of interest (e.g., the first impacted or the most susceptible), providing extensions to the upstream and downstream links. Ontology merging and alignment tools required have been adopted from [13] and [19].

The network is thinned out (step 9) according to the information-theoretical algorithm of Cheng *et al*. [20], although more advanced approaches can be considered as those using maximum likelihood estimation (e.g., see [21]. The algorithm has three subsequent phases termed drafting, thickening and thinning. In the drafting phase, the algorithm establishes, from the data, the mutual information for each pair of variables and constructs a draft digraph from this information. In the thickening phase, the algorithm adds arcs between pairs of nodes if the corresponding variables are not conditionally independent given a certain conditioning set of variables. In the thinning phase, to conclude, each arc of the graph obtained so far is examined using conditional independence tests, and is removed if the two variables connected by the arc prove to be conditionally independent.

3   Implementation

3.1 Modelling denitrification processes in rivers

The construction of a suitable physical model for a river with clearly defined boundaries should take under consideration many parameters and, inevitably, it should rely on many assumptions. A river stretch may be defined by the area between two inflowing tributaries. A major difficulty arises in drawing borderlines towards a lake or an estuarine or adjacent coastal areas; if such borderlines are drawn arbitrarily, one would get arbitrary volumes, areas and mean depths and the mass-balance model would lose predictive power (see, e.g., [22]). A suitable physical model for representing nitrification/denitrifications processes is the river mass-balance model handling internal fluxes (fig. 3), modified from the suspended particulate matter (SPM) and phosphorus model presented in [23]. The model is simple to apply in practice since all driving variables may be readily accessed from maps and standard monitoring programs; it contains with four compartments: (I) bank area, with sub-compartments for inflow (dry land) and outflow (we land), (II) fixation, (III) transport/ accumulation, and (IV) the upstream river stretch.
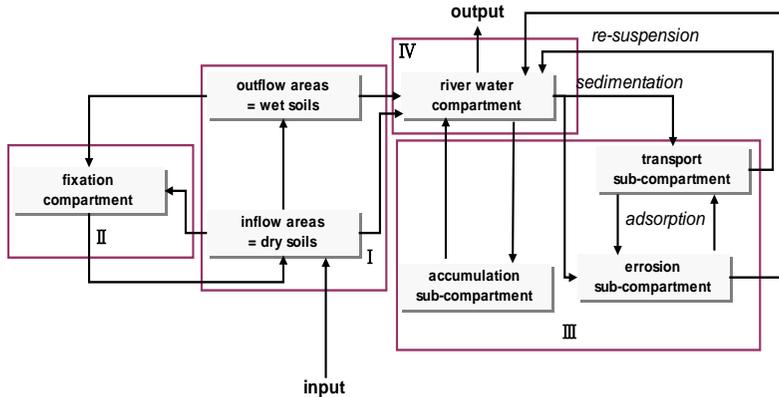
Figure 3:   The river model adopted herein for modelling nitrification and denitrification processes.

The river nitrogen load (step 6) is often of serious concern because of its potential to cause adverse effects in the receiving water systems [24]. Among the various nitrogen groups, the dissolved inorganic nitrogen species, especially ammonium, nitrite and nitrate species, have the greatest impact on aquatic biotic and abiotic constituents, because, owing to their high biovailability, they are fastly uptaken by microorganisms [24]. In the fixation compartment, ammonium tends to be oxidized to nitrate in a two-step process ($[NH_4^+] \rightarrow [NO_2^-] \rightarrow [NO_3^-]$) by aerobic chemoautotrophic bacteria (*Nitrosomonas* and *Nitrobacter*, primarily); the process can even occur at very low oxygen levels [25]. NH_4^+, NO_2^- and NO_3^- may however be removed from water by macrophytes, algae and bacteria which assimilate them as sources of nitrogen [24, 25].

The biological denitrification mechanism makes use of nitrate as the terminal electron acceptor in low-oxygen environments [26]. In this process, denitrifying bacteria decrease inorganic nitrogen into innocuous fundamental nitrogen gas. Denitrification can only take place in the anoxic parts of the system (i.e., at the accumulation sub-compartment near the water-sediment interface), as the presence of dissolved oxygen suppresses the enzyme systems required for this process [25]. High concentrations of nitrate entering this compartment can lead to more vigorous and robust populations of denitrifiers at the sediment [24]. The rate of denitrification is influenced by many factors, including nitrate concentration, microbial flora, type and quality of organic carbon source, hydroperiods, different plant species residues, the absence of O_2, redox potential, moisture, temperature (and especially its variations), pH values, presence of denitrifiers, soil type, and the presence of overlying water [25].

In view of the above, the water quality control variables and parameters identified in the river ecosystem (step 6) include: dissolved oxygen (DO), phytoplankton as carbon (PHYT), carbonaceous biochemical oxygen demand (CBOD), ammonium nitrogen (NH_4^+), nitrate and nitrite nitrogen (NO_3), ortho-phosphorus or inorganic phosphorus (OPO_4), organic phosphorous (OP). The

ontological network developed at step 7 describes a basic transformation process including photosynthesis, uptake, respiration, nitrification, denitrification, benthic flux, sediment suspension, and external loads, putting emphasis on the water column and the sediment anaerobic layer at the accumulation sub-compartment. After thinning (step 9), the nodes with the higher in- and out-degrees (step 10) are shown in fig. 4 in descending order of significance; boxed species are the nodes with the higher in- and out-degrees, whereas circled numbers indicate the strategic nodes. These nodes can be used to verify the holistic model since monitoring and parameter-value estimation is feasible.
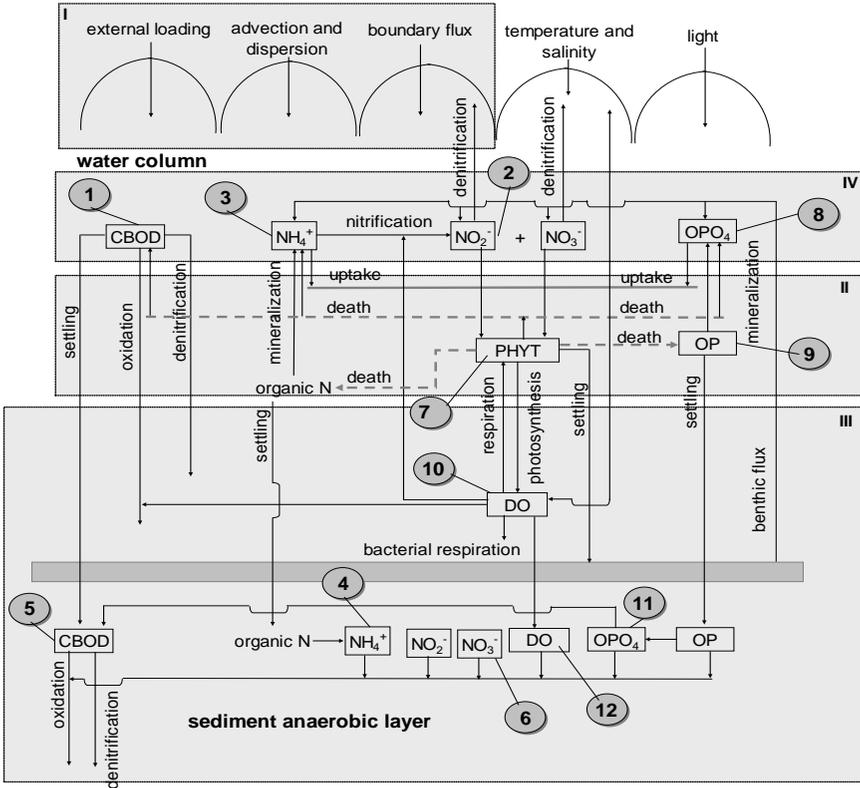
Figure 4:   The ontological network (after thinning) describing river nitrification/denitrification processes in a vertical section of the physical model.

3.2 Oil weathering at sea

A large number of oil spill weathering models are in use today. These range in capability from simple trajectory, or particle-tracking models, to three-dimensional trajectory and fate models that include simulation of response actions and estimation of biological effects. Models frequently fail [27] due to

the close interdependence of oil spill weathering processes. These interrelations have been successfully handled by the ontological network constructed herein, which provides the relevant partonomic relations to describe a number of processes occurring simultaneously within a micro- or macro-region of the oil slick: inputs, fates, and effects. As a case example, we consider the fate of polycyclic aromatic hydrocarbons (PAHs) contained in the oil slick. Oxidation of petroleum constituents into the sea (especially by the microbes present) may, also, result in PAH formation, necessitating the estimation of kinetic parameters for both directions. The knowledge representation in Fig. 5 uses two levels of knowledge (surface and deeper) that communicate through a middle level that deals with their inter-relations; a higher granularity level is also available to help estimate metrological parameters. These inter-relationships among the physical, chemical, and biological processes that crude oil is subjected to upon its introduction into the marine environment, lead to multi-scale weathering and its final transportation away from the incidence site.
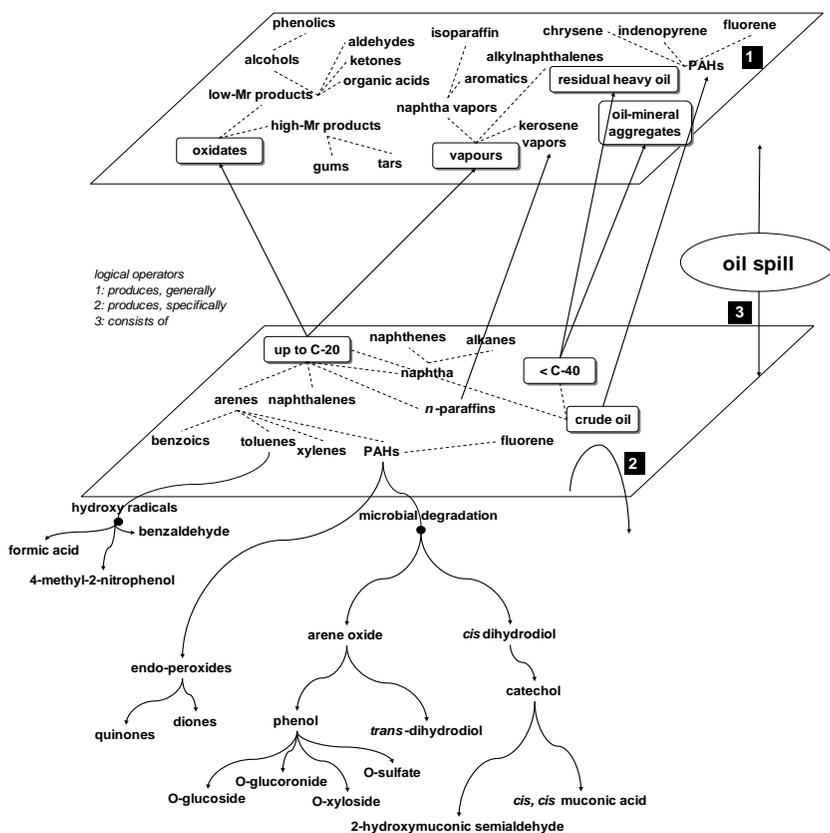
Figure 5:   Schematic representation of the oil weathering ontology.

The fate and persistence of oil in water are controlled by processes that vary considerably in space and time; thus, it is not possible to derive simple generic relationships between petroleum mass loadings and ambient concentrations. For example, the addition of dispersants may change significantly degradation kinetics (step 4), increasing the uncertainty on the accuracy of nodes ranking in step 7. Since modelling in such cases is taken to support environmental management activities, extended measurements are required for validation in order to enhance the reliability of restoration monitoring.

4   Discussion and conclusions

The use of the ontological platform is extremely valuable in outlining/elucidating processes and determining control variables in aquatic ecosystems. The high complexity of ecological processes and their inter-relations might (and commonly do) conceal the significant links in mass/energy and other interdependence chains that, actually, drive a constantly changing system responding to seasonal variations and pollution load alterations. Moreover, the profound non-linearity of the quantitative relations that govern the interdependence chains create the pre-requisites for chaotic behaviours that could result in irreversible adverse effects on the ecosystem under consideration, as well as the wider environment.

In the present work, ecosystem-relevant spatio-temporal parameters at various scales are revealed and linked to support dynamic multi-scale modelling. This scheme enables knowledge to be used not only for representation but also for reasoning at conceptual level, proven to be useful in real-world problem solving; that requires, inevitably, the involvement of domain experts, herein realized by integrating the modelling construction process within a framework of 2nd order cybernetics.

The implementation of the proposed framework for holistic modelling is expected to aid decision making in a diverse range of water management activities, including: predictive modelling for wastewater regulation, water allocation decisions and infrastructure operations, marine pollution monitoring (especially at ports where the low depths facility wide dispersions), development of ambient guidelines and in-stream targets to support water management planning, environmental performance measurement and state of the environment reporting, negotiation of trans-boundary water management agreements, as well as research into a wide variety of issues including aquatic ecosystem contamination, use, impairment, and restoration.

Acknowledgement

References

[1]   Knoke, T. and Seifert, T., Integrating selected ecological effects of mixed European beech–Norway spruce stands in bioeconomic modelling, *Ecological Modelling*, **210**, pp. 487-498, 2008.

[2]   Sillero, N., What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, **222**, pp. 1343-1346, 2011.

[3]   Villa F., Integrating modelling architecture: a declarative framework for multi-paradigm, multi-scale ecological modelling. *Ecological Modelling*, **137**, pp. 23-42, 2001.

[4]   Ingram, G.D., Cameron, I.T. and Hangos, K.M., Classification and analysis of integrating frameworks in multiscale modelling. *Chemical Engineering Science*, **59**, pp. 2171-2187, 2004.

[5]   Siontorou, C.G. and Batzias, F.A., Error identification/propagation/ remediation in biomonitoring surveys – A knowledge-based approach towards standardization via fault tree analysis. *Ecological Indicators*, **11**, pp. 564-581, 2011.

[6]   Hesse, C., Krysanova, V., Päzolt, J. and Hattermann, F.F., Eco-hydrological modelling in a highly regulated lowland catchment to find measures for improving water quality, *Ecological Modelling*, **218**, pp. 135-148, 2008.

[7]   Marmion, M., Hjort, J., Thuiller, W. and Luoto, M., Statistical consensus methods for improving predictive geomorphology maps, *Computers & Geosciences*, **35**, pp. 615-625, 2009.

[8]   Naveh, Z., Allen, E., Laszlo, E. and Antrop, M., *Transdisciplinary Challenges in Landscape Ecology and Restoration Ecology* – An Anthology, Springer, pp. 353-402, 2007.

[9]   Liu, Y., Gupta, H., Springer, E. and Wagener, T., Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management, *Environmental Modelling & Software*, **23**, pp. 846-858, 2008.

[10] Batzias, F.A and Siontorou C.G., Creating a specific domain ontology for supporting R&D in the science-based sector – The case of biosensors, *Expert Systems with Applications*, **39**, pp. 9994-10015, 2012.

[11] Darlington, M.J. and Culley, S.J., Investigating ontology development for engineering design support, *Advanced Engineering Informatics,* **22**, pp. 112-134, 2008.

[12] Malhorta, R., Meta-modelling framework: A new approach to manage meta-modelbase and modeling knowledge, *Knowledge-Based Systems,* **21**, pp. 6-37, 2008.

[13] Kotis, K., Vouros, G.A. and Stergiou, S., Towards automatic merging of domain ontologies: the HCONE-merge approach, *Web Semantics: Science, Services and Agents on the World Wide Web*, **4**, pp. 60-79, 2006.

[14] Bermudez, L., Graybeal, J.B., Anthony W.I., Lowry, R. and Dawn, W., Construction of marine vocabularies in the Marine Metadata

Interoperability Project, *Proceedings of MTS/IEEE OCEANS 2005*, **3**, pp. 2578-2585, 2005.

[15] Li, H., Mynett, A., Penning, E. and Qi, H., Revealing spatial pattern dynamics in aquatic ecosystem modelling with Multi-Agent Systems in Lake Veluwe, *Ecological Informatics*, **5**, pp. 97-107, 2010.

[16] Chau, K.W., An ontology-based knowledge management system for flow and water quality modelling, *Advances in Engineering Software*, **38**, pp. 172-181, 2007.

[17] Islam, A.S. and Piasecki, M., Ontology based web simulation system for hydrodynamic modelling, *Simulation Modelling Practice and Theory*, **16**, pp. 754-767, 2008.

[18] ASTM E2107 – 06 Standard Practice for Environmental Regulatory Compliance Audits, Book of Standards Volume: 11.05.

[19] Seddiqui, Md.H. and Aono, M., An efficient and scalable algorithm for segmented alignment of ontologies of arbitrary size, *Web Semantics: Science, Services and Agents on the World Wide Web,* **7**, pp. 344-356, 2009.

[20] Cheng, J., Bell, D., and Liu, W., Learning Bayesian networks from data: an efficient approach based on information theory. *Proceeding of the Sixth ACM International Conference on Information and Knowledge Management*, pp. 325–331, 1997.

[21] Hastie, T., Tibshirani, R. and Friedman, J., *The elements of statistical learning: data mining, inference, and prediction,* Springer, New York, 2001.

[22] Håkanson, L., A new general mechanistic river model for radionuclides from single pulse fallouts which can be run by readily accessible driving variables, *Journal of Environmental Radioactivity*, **80**, pp. 357-382, 2005.

[23] Håkanson, L., A new general dynamic model predicting radionuclide concentrations and fluxes in coastal areas from readily accessible driving variables, *Journal of Environmental Radioactivity,* **78**, pp. 217–245, 2005.

[24] Bastviken, S.K., Eriksson, P.G., Martins, I., Neto, J.M., Leonardson, L. and Tonderski, K., Potential nitrification and denitrification on different surfaces in a constructed treatment wetland, *Journal of Environmental Quality,* **32**, pp. 2414-2420, 2003.

[25] Rabalais, N.N., Nitrogen in aquatic ecosystems, *Ambio* **31**, pp. 102–112, 2002.

[26] Reddy, K.R. and Patrick, W.H., Nitrogen transformations and loss in flooded soils and sediments, *Critical Reviews in Environmental Control,* **13**, pp. 273–309, 1984.

[27] Reed, M., Johansen, O., Brandvik, P.J., Daling, P., Lewis, A., Fiocco, R., Mackay, D. and Prentki, R., Oil spill modeling towards the close of the 20th century: Overview of the state of the art, *Spill Science & Technology Bulletin*, **5**, pp. 3-16, 1999.