



Article

# An Integrated Software Framework to Support Semantic Modeling and Reasoning of Spatiotemporal Change of Geographical Objects: A Use Case of Land Use and Land Cover Change Study

Wenwen Li 1,\*, Xiran Zhou 1 and Sheng Wu 1,2

- School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287-5302, USA; xrzhou@asu.edu (X.Z.); wushengcq@126.com (S.W.)
- School of Computer and Information Science, Southwest University, Chongqing 400715, China
- \* Correspondence: Wenwen@asu.edu; Tel.: +1-480-727-5987

Academic Editor: Wolfgang Kainz

Received: 23 June 2016; Accepted: 20 September 2016; Published: 30 September 2016

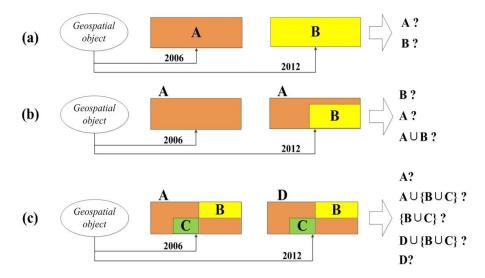
Abstract: Evolving Earth observation and change detection techniques enable the automatic identification of Land Use and Land Cover Change (LULCC) over a large extent from massive amounts of remote sensing data. It at the same time poses a major challenge in effective organization, representation and modeling of such information. This study proposes and implements an integrated computational framework to support the modeling, semantic and spatial reasoning of change information with regard to space, time and topology. We first proposed a conceptual model to formally represent the spatiotemporal variation of change data, which is essential knowledge to support various environmental and social studies, such as deforestation and urbanization studies. Then, a spatial ontology was created to encode these semantic spatiotemporal data in a machine-understandable format. Based on the knowledge defined in the ontology and related reasoning rules, a semantic platform was developed to support the semantic query and change trajectory reasoning of areas with LULCC. This semantic platform is innovative, as it integrates semantic and spatial reasoning into a coherent computational and operational software framework to support automated semantic analysis of time series data that can go beyond LULC datasets. In addition, this system scales well as the amount of data increases, validated by a number of experimental results. This work contributes significantly to both the geospatial Semantic Web and GIScience communities in terms of the establishment of the (web-based) semantic platform for collaborative question answering and decision-making.

**Keywords:** geospatial semantic modeling; change detection; land use and land cover; semantic reasoning; spatial reasoning

#### 1. Introduction

The advancement of remote sensing techniques and the Global Earth Observation System of Systems (GEOSS) makes it possible to collect vast amounts of environmental data about the Earth's surface at a finer resolution. Change detection of land use land cover (LULC) is one of the key research areas of remote sensing change detection [1–4]. Theories and technologies for change detection have been applied to a wide range of geospatial applications, including landscape monitoring [5,6], natural resource estimation [7] and environmental management [8]. Although several techniques [9] can accurately detect LULC changes from remote sensing imagery, there exist great challenges to effectively organize and model such information from multi-temporal remote sensing images.

These challenges stem from the fact that change in one or more dimensions can cause a spatial object to split into multiple objects, merge with another spatial object(s) or dissolve into a new object. Furthermore, changes that combine multiple relationships (dimensions) such as these are very difficult to model. Figure 1 illustrates three challenges facing any semantic model for change information modeling from time series remote sensing data.



**Figure 1.** Illustration of the challenges of semantic concept modeling for Land Use and Land Cover Change (LULCC). (a) Represents a change of LULC for the entire region; (b) represents local LULC change of a region; (c) represents a global change of LULC in the region while local LULC remains the same.

In the first case (Figure 1a), we observe the entire geospatial object's LULC changing over time. The geospatial object demonstrates two distinct LULC attributes at different timestamps: A in 2006 and B in 2012. Semantic models need to integrate both definitions of this object. For the second case (Figure 1b), local attributes change within an otherwise globally-uniform object. The geospatial object contains only one attribute A in 2006 while maintaining most of that attribute with partial attribute B changes in 2012. When constructing a semantic model for the attributes of this geospatial object, it will be difficult to categorize, as it could plausibly be described as A, B or even A and B simultaneously (the case in Figure 1b). In the third case (Figure 1c), the global attribute of the object may change while local attributes remain the same. In 2006, the geospatial object has a global attribute A and contains two subsets B and C. In 2012, the global attribute A changes to the global attribute D while the local attributes C and B remain unchanged. During semantic modeling for this geospatial object, the categories become much more complex as the geospatial object may be described as A, D, or "B and C", or as an even more complicated expression, such as A with subsets of B and C or D with subsets of B and C.

To address these challenges, comprehensive spatio-temporal change detection research should consist of detecting changes (where they occur), semantically defining the changes (what are they), effectively tracking the changes (how they evolve) and connecting these changes with other geophysical or human processes (how they happen). This requires a new modeling approach that can store and capture all of these dimensions of information [10]. Therefore, change detection requires a multi-dimensional data model that includes spatial, spectral, temporal and attribute dimensions. Due to the dynamics of LULC Change (LULCC), semantic reasoning is also needed to be able to trace the evolution of an area and explore its spatial, temporal and topological relationships with nearby areas [11]. Finally, it is critical that LULCC representation integrates knowledge from other sources, such as land and resource planning information, beyond that extracted from the remote sensing imagery [12].

To model this complexity and effectively manage LULCC information from time series geospatial data (remote sensing imagery), we propose and implement an integrated computational framework to support the modeling, semantic and spatial reasoning of change information with regard to space, time and topology. The remainder of the paper is organized as follows: Section 2 reviews the relevant work in LULC and spatiotemporal data modeling and discusses the remaining issues in these area. Section 3 proposes the conceptual model for LULCC modeling, its ontological implementation, the definition of queries and rules to infer new knowledge from existing data, as well as the implementation detail of the semantic platform for spatio-temporal and semantic reasoning. Section 4 demonstrates the prototype with an illustration of several cases of change query and change trajectory tracing. Section 5 validates the scalability of the proposed system through a number of experiments. Section 6 concludes the work, highlights our contributions and discusses future research directions.

#### 2. Literature Review

There have been many attempts to address the challenges in the formal organization of change detection results from remote sensing imagery [13–15]. The challenges belong to two broader categories: (1) how to address the vagueness or semantic ambiguity in the definition of an LULC type; and (2) how to semantically model the change information regarding its space, time and context from time series remote sensing data.

# 2.1. Progress in Enhancing the Semantic Interoperability of Existing LULC Classification Systems

The first challenge is a well-known semantic interoperability problem. There exists a large number of classification standards developed by different organizations used in various disciplines for different purposes and at different levels of abstraction [16]. For instance, widely-applied LULC classification systems alone include the (Modified) Anderson LULC Classification System (LCCS) (Anderson LULC) [17], CLC (Corine Land Cover) [18], Land-Based Classification Standards (LBCS) [19], FAO LCCS [20], LULC Classification of United Kingdom National Land Use Database (NLUD) [21], among many others. These different classification systems define LULC classes differently, impeding the exchange, integration and reuse of data. These semantic issues also make identifying and interpreting change processes more difficult [22].

Efforts to overcome this semantic heterogeneity include the development of a homogeneous classification standard to apply to existing LULC datasets from multiple sources. For example, INSPIRE [23] is a European Union initiative to create a spatial data infrastructure to harmonize geospatial information and provide one-stop access to datasets from various European countries. Uniform classification systems and metadata standards will be applied such that the heterogeneity issue can be resolved. Strategies of this kind can be considered as top-down approaches, as all member states or countries will use consensus classification standards from an authority to improve mutual understanding of the data.

Other efforts include a bottom-up approach that is dedicated to developing a crosswalk among different LULC classification systems through semantic analysis. This strategy is also known as ontology matching or alignment and requires the calculation of the semantic similarity between different LULC class definitions. Typical methods include geometric [24], feature-based [25,26], network [27], alignment [28] and information-theoretic [29] modeling. Because a land type definition is usually given in a text block, text analysis techniques were also explored for extracting the spatial and aspatial attributes of an LULC class to enrich the formal definition in the LULC ontology [30]. Although these works have substantially improved the semantic interoperability among existing LULC classification systems through ontological formalization, they focus mostly on matching the conceptual definition of land use types, rather than specific spatial objects identified from remote sensing images, which contain more information than a single LULC type. The semantic modeling of the spatiotemporal change of a particular region, however, is understudied. One of this paper's major goals is to solve this issue.

## 2.2. Progress in Developing a Spatiotemporal Reasoning Platform

In comparison with the widely-studied semantic interoperability problem in the LULC classification systems, fewer studies [31–37] on the semantic modeling of change information regarding its spatial, temporal and other attributes, and especially on integrating these semantic models into an operational spatial reasoning framework, are found. This is partially due to the complexity in modeling spatial objects, which contain four distinct dimensions: (1) geographical (i.e., location); (2) geometric dimension (i.e., shape or size); (3) topological (i.e., intersections with other spatial objects); and (4) temporal, (i.e., creation and cessation time). Beyond this, modeling complexity also increases dramatically when a change process involves complex relationships combining multiple dimensions, such as urbanization, that need to be captured.

Some pioneering works in this realm include those by Hornsby and Egenhofer [31], Worboys [32], Claramunt and Jiang [33], Jiang and Worboys [34] and others, which lay the theoretical foundation for representing the change of a spatial object. In practice, Varanka and Usery [35] developed a surface water ontology to formally define the United States Geological Survey (USGS) National Hydrological Dataset. In this work, a spatial extension is made to general ontology, which focuses only on conceptual naming [36], to model hydrological features. The spatial property of a water body feature is annotated by its spatial extent in the ontology, and a temporal quality property is used to describe the characteristics (i.e., ephemeral, intermittent etc.) of its water flow. Batsakis and Petrakis [37] developed a 4D-fluent model to express the existence of an entity that relates to some defined time interval. This is the most well-known model that can capture the dynamics of an entity in an ontology. However, this model increases the complexity for querying the temporal changes and is limited in maintaining a relationship between geometry and semantics, thereby not suitable for modeling land use change data, of which the location/geometry is an indispensable component.

Very recently, Arenas and colleagues [9] developed a spatiotemporal data model, LC3 (Land Cover 3), to study LULCC. Several change processes, including splitting, merging, separation and annexation, are defined. This model has a clear advantage in modeling change information; however, all of the topological relationships, such as "contain" and "within," are predefined in the ontology. This causes issues in the modeling and querying of large numbers of spatial objects, the spatial relationships among which are the exponent of the total number of objects.

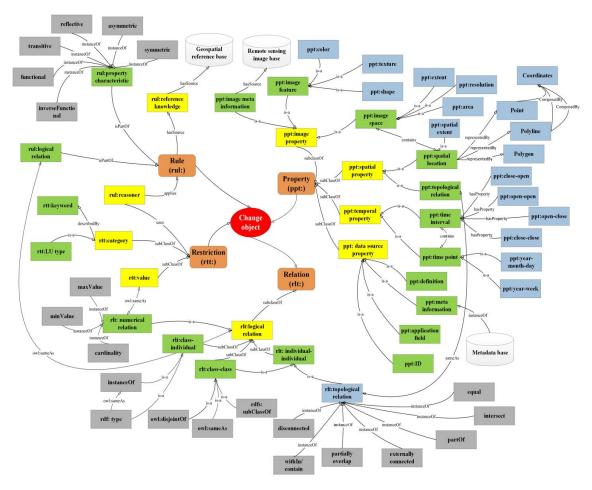
In the W3C (World Wide Web Consortium) community, there has been research that extends the ontological modeling languages, RDF (Resource Description Framework) [38] or OWL (Web Ontology Language), to integrate the time dimension [39]. New specifications, such as OWL-Time, have been developed [40,41]. Some space-time modeling works applying these standards can be found in [42–45]. For instance, Bereta and colleagues developed a new query language stSPARQL (space-time semantic web query language) to support expressive encoding of time in OWL ontologies [43]. Huang and Deng [44] pointed out the importance of enabling spatio-temporal reasoning based on geo-ontology and SWRL (Semantic Web Rule Language); however, they only touched very briefly and conceptually on the solution. Batsakis and Petrakis [45] used OWL 2.0 and SWRL to define spatio-temporal information, mainly focusing on encoding the fuzzy spatial relations on directions (i.e., left-of) of spatial objects. The attributes, such as land use properties and change information, are not modeled.

To overcome the aforementioned limitations in LULC and spatiotemporal semantic modeling works, this paper tends to develop an operational software framework to enable the dynamic reasoning of the spatiotemporal change of geospatial objects within the context of modeling and querying about LULC change. The next section describes the construction of the LULCC ontology, the definition of the semantic rules to support reasoning and the implementation of the software framework, which operates on the knowledge in the ontology and reasoning rules for building the dynamic semantic platform.

## 3. Building Blocks of the Semantic Platform

# 3.1. Geospatial Ontology for LULCC

The proposed geospatial ontology represents a spatial object, its static (i.e., property of the image from which a geospatial object is obtained) and dynamic properties (i.e., change information over time) and its relationships with other nearby spatial objects. It formally represents LULCC, providing a semantic relationship model. Built upon change annotations, this LULCC geospatial ontology uses four core components to describe the change objects: property, relation, rule and restriction. They are shown in orange in Figure 2 as the first tier of the ontology.



**Figure 2.** A geospatial ontology for semantic modeling of change information. Nodes of the same color refer to subdivisions at the same level.

In the ontology, different entities are assigned with different prefixes to refer to the facet or namespace to which they belong. For example, ppt:spatial\_property, ppt:image\_property, ppt:temporal\_property and ppt:thematic\_property indicate the spatial, image, temporal and thematic properties of a spatial object. Beyond the four core components, entities in yellow and green represent the first-level and second-level subclasses, respectively. For instance, "property" is further divided into "spatial property", "temporal property", "image property" and "data source property", describing the spatial and temporal property of a spatial object, as well as the characteristics of the remote sensing image from which the object is extracted. The "rule" component describes the inference rules in a machine-understandable format, and the rules are categorized into a reflective rule or a symmetric rule to support semantic reasoning. Blue nodes represent the type of an instance. For instance, the "polygon" class is a type of "spatial location" class to describe the spatial property of a feature with

an area property, such as a parcel. The leaf nodes in grey in the "relation" subtree are subtypes of instances of its parent. They will be implemented as the predicates in the triple structure <subject, predicate, object> in OWL. External data for this ontological model come from the reference, remote sensing image and metadata databases.

## 3.2. Defining Reasoning Rules

Rules provide support for logical reasoning based on reference knowledge from diverse sources. Reference knowledge is the experts' professional understanding extracted from the geospatial reference base and offers the basis for reasoning. Our semantic model supports two types of reasoning for detecting spatiotemporal change: logical reasoning and spatial reasoning. Logical reasoning utilizes the characteristics of an OWL property of being transitive or reflective to infer spatial knowledge. It also utilizes rules specifically defined for making inference on the change of land use within a geographical area or of a spatial object.

Rule 1: For any two objects  $x_1$  and  $y_1$ , let it be known that  $x_1$  exists at time point  $t_1$  and  $y_1$  exists at time point  $t_2$  ( $t_1 < t_2$ ). The spatial boundary of object  $x_1$  isSameAs the spatial boundary of object  $y_1$ , but the LULC of object  $x_1$  isNotSameAs the LULC of  $y_1$ ; we can infer that LULC changes happened between the two timestamps. Note that relations, such as "isSameAs" and "isNotSameAs", are predefined properties in the LULCC ontology. The rule can be written as:

$$\left. \begin{array}{c} \textit{hasSpatialBoundary} \ (?x_1,?l_1) \wedge \textit{hasSpatialBoundary} \ (?y_1,?l_2) \wedge \\ \textit{hasLULC} \ (?x_1,?z_1) \wedge \textit{hasLULC} \ (?y_1,?z_2) \wedge \\ \textit{isSameAs} \ (?l_1,?l_2) \wedge \\ \textit{isNotSameAs} \ (?z_1,?z_2) \end{array} \right\} \Longrightarrow \textit{LULCC} \ (?x_1,?y_1) \qquad (1)$$

where  $l_1$  and  $l_2$ , and  $z_1$  and  $z_2$  represent the spatial boundaries and LULC types for objects  $x_1$  and  $y_1$ . is LULCC ( $?x_1, ?y_1$ ) represents the function to check whether there is some LULC change for an object.

Rule 2: Given two objects  $x_1$  and  $x_2$ ,  $x_1$  exists at time point  $t_1$ , and  $y_1$  exists at time point  $t_2$  ( $t_1 < t_2$ ). Assume  $x_1$  has a spatial boundary of  $l_1$ , and  $y_1$ 's spatial boundary is  $l_2$ . If  $l_1$  is completely within  $l_2$ , then a merge process occurs from  $t_1$  to  $t_2$ , since  $y_1$  must come from the merge among  $x_1$  and other objects. The rules to determine a merge process can be formalized in Expression (2):

$$\left. \begin{array}{c} hasSpatialBoundary\left(?x_{1},?l_{1}\right) \wedge hasSpatialBoundary\left(?x_{2},?l_{2}\right) \wedge \\ & within\left(l_{1},l_{2}\right) \wedge \\ & isNotSameAs\left(?l_{1},?l_{2}\right) \end{array} \right\} \Longrightarrow mergeInto\left(?x_{1},?x_{2}\right) \quad \ \left\{ \begin{array}{c} \\ \\ \end{array} \right\}$$

Rule 3: Given an object  $x_1$ , which exists at time  $t_1$ , and  $y_i$   $\{i \ge 1\}$  is any object that exists at  $t_2$   $(t_1 < t_2)$ , assume  $x_1$  has a spatial boundary of  $l_1$ , and  $y_i$ 's spatial boundary is  $l_{y_i}$ ; a split process can be determined if  $x_1$  intersects with more than  $y_i$  at time  $t_2$ . The rules to determine a split process can be formalized in Expression (3):

The above rules cover the basic reasoning logic to infer whether a LULCC change has occurred between two timestamps in terms of attribute changes (Rule 1) and spatial changes (Rules 2 and 3). These rules can be combined to generate complex rules, which model both spatial and temporal changes. The execution of the reasoning processes (using SWRL) requires the integration of both spatial and logic reasoning, as well as the support from the backbone LULCC ontology.

The next section describes the implementation of a spatial-enabled semantic query and reasoning engine in detail.

## 3.3. An Integrated Software Framework to Support Semantic Query and Reasoning

To support the semantic query and reasoning about LULC change in an automatic and intelligent manner, the ontology or knowledge base must be established in a machine-understandable way. Therefore, a semantic query and reasoning framework was built to support the dynamic query and reasoning of the change data, both spatially and non-spatially. Figure 3 depicts the three-layer software framework to accomplish this.

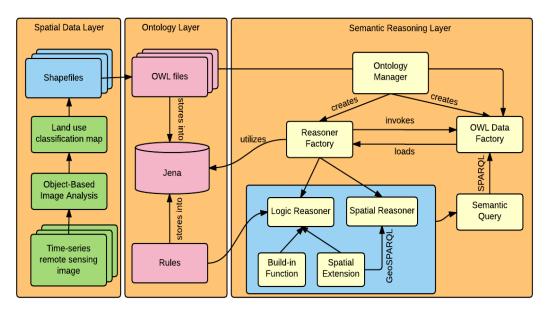


Figure 3. An integrated software framework for LULCC query and reasoning.

The layer on the left is the spatial data layer. Time series remote sensing images are preprocessed using Object-Based Image Analysis (OBIA) techniques to generate the LULC classification maps. In detail, eCognition was used to segment each satellite image at multiple scales (a higher scale and a lower scale). Based on the result obtained from lower-scale segmentation, OBIA classification was further conducted by eCognition. Several commonly-used low-level image features, including GLCM (Gray-Level Co-occurrence Matrix) homogeneity, GLCM contrast, GLCM dissimilarity, GLCM entropy, as well as the mean and standard deviation of segmented image regions, were selected for classification. After receiving the original classification result generated from eCognition, the errors were manually corrected based on the results from a higher-scale segmentation. Note that since the focus of this work is semantic modeling and reasoning, manual work was introduced in the classification process to remove uncertainties and ensure accurate classification results.

The product generated from the above procedure is a raster map, and it is then vectorized to generate a polygonal file, which denotes the subdivisions of the study area according to their LULC types. These vector data are compiled into ESRI shapefiles or imported into spatial databases, such as PostGreSQL (SQL: Structured Query Language). The dataset is then composed of two parts, the spatial part that records the geographical boundary of each polygonal object and the alphanumerical part that records the attributes, such as LULC type, year of existence of a spatial object and related information integrated from other knowledge bases.

Next, the LULC data are loaded from the backend data repository or shapefiles using the GeoTools Feature API [46]. As indicated in the ontological schema defined in Figure 2, data records are serialized into triples and written into the native triple store of Jena by applying the Jena API [47]. The logical rules defined in Section 3.2 are also imported into the Jena data store. This flow of data preparation in

a machine-understandable way provides strong support to the reasoning procedure depicted in the "semantic reasoning layer".

Initially, an ontology manager, which operates and coordinates all data and reasoning modules, is constructed. This manager directly commands two components: the reasoner factory and the OWL data factory. The OWL data factory is responsible for loading semantic data from the Jena triple store to support various semantic queries in standard SPARQL (Semantic Web Query Language) [48] format. This component supports only the queries of static data in the triple store. In other words, only data and knowledge that already exist in the ontology can be returned. Providing answers to dynamic queries, where the knowledge is previously unknown, requires support from the reasoner factory component.

As discussed, successful reasoning about the spatial and temporal change of a land use object not only relies on powerful logic reasoning capability, but also requires the seamless integration with the spatial reasoning process. Consider Rule 3, the split process in Section 3.2, as an example. One of the conditions to support the detection of whether an object  $x_1$  is split into multiple objects in the next timestamp  $t_2$  is to take the intersection of  $x_1$ 's boundary with all objects existing at  $t_2$ . This "intersect" function is not part of the built-in reasoning function in Jena; therefore, a spatial extension needs to be implemented within the reasoning factory, and a GeoSPARQL query interface is enabled to support the communication between the spatial reasoner and the spatial extension module.

By spatially enabling existing semantic reasoning capability, this integrated framework provides comprehensive support to LULCC query and reasoning over space and time. In the next section, we describe the application of the above framework using time series remote sensing data in Ningbo, China.

#### 4. Demonstrations and Discussions

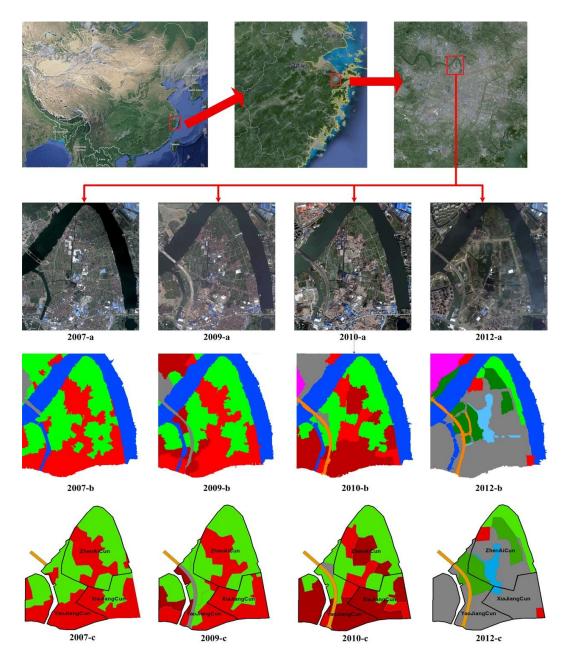
### 4.1. Data and Study Area

The study area considering in this work was the Wantou community of Ningbo, in Zhejiang Province, China. There are two reasons we chose Wantou. First, Wantou has undergone fast development. Its historical landscape, current conditions and planning are of great interest to local government. Second, although Wantou is a village-style community in Ningbo City primarily composed of residential areas, it has a great potential in landscape planning because of the predominant natural environment and therefore has undergone substantial reconstruction in modern times. Thus, a significant spatiotemporal change occurred from 2007 to now, making it a good case study for modeling LULCC.

This study uses four high-resolution remote sensing images (QuickBird) taken at four different times (Figure 4). This imagery includes the landscape condition of Wantou Community in 2007, 2009, 2010 and 2012. Table 1 lists the image dates and the band characteristics. In order to build a geospatial semantic model for change information extracted from remote sensing imagery, we classify and annotate these four images using the OBIA method and then use GIS software (i.e., ESRI ArcMap) to convert the classification results from raster data to an object-based vector format.

Satellite Image	QuickBird
	22 August 2007
Date	5 June 2009
	12 March 2010
	16 April 2012
	Blue: 450–520 nm
Waveband	Green: 520-600 nm
	Red: 630-690 nm
Spatial resolution	0.65 m

Table 1. Detail information about the QuickBird images in use.



**Figure 4.** Demonstration of remote sensing image samples (**a**), image classification results (**b**) and GIS generation results (**c**).

Wantou consists of three villages, ZhenAiCun (Zhenai Village), XiaJiangCun (Xiajiang Village) and YaoJiangCun (Yaojiang Village), and the Wantou Bridge. In Figure 4, 2007-a, 2009-a, 2010-a and 2012-a show the original QuickBird images taken in 2007, 2009, 2010 and 2012. 2007-b, 2009-b, 2010-b and 2012-b show the classification results of the four original images using the aforementioned OBIA approach. 2007-c, 2009-c, 2010-c and 2012-c depict the GIS-generated classification results. The land use category is displayed in different colors. Red is residential land, brown, vacant land, light green, farmland, dark green, green land, orange, transportation avenues, grey, construction sites, and blue, lakes. The land use composition of Wantou's sub-components from 2007 to 2012 are shown in Table 2.

Village	Land Cover (2007)	Land Cover (2009)	Land Cover (2010)	Land Cover (2012)
ZhenAiCun	Residential (34.5%) Agriculture (65.5%)	Residential (34.5%) Agriculture (65.5%)	Residential (64.0%) Agriculture (16.7%) Urban vacant (19.3%)	Residential (4.6%) Agriculture (20.8%) Under construction (37.6%) Park (25.2%) Waters (11.8%)
XiaJiangCun	Residential (47.95%) Agriculture (31.22%) Industrial (20.83%)	Residential (47.95%) Agriculture (31.22%) Industrial (20.83%)	Residential (30.1%) Agriculture (26.5%) Urban vacant (11.3%) Industrial (32.2%)	Industrial (4.1%) Under construction (92.1%) Waters (3.7%)
YaoJiangCun	Residential (36.9%) Agriculture (44.6%) Industrial (18.5%)	Residential (29.2%) Agriculture (35.4%) Under construction (20.8%) Industrial (14.6%)	Residential (9.1%) Agriculture (22.5%) Urban vacant (44.8%) Under construction (3.1%) Transportation (7.8%) Industrial (12.7%)	Residential (22.8%) Industrial (3.0%) Under construction (49.2%) Park (17.3%) Transportation (7.8%)
Wantou Bridge	Under construction (100%)	Transportation (100%)	Transportation (100%)	Transportation (100%)

Table 2. Land use composition of Wantou community (2.45 km<sup>2</sup> in area) from 2007 to 2012.

## 4.2. Ontological Implementation for the LULCC Data

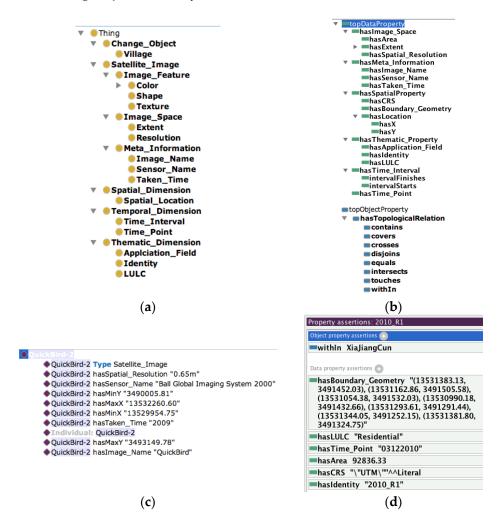
For change detection computation, we obtain all change objects with land use annotations in Wantou Community in 2007, 2009, 2010 and 2012. We first create a geospatial ontology according to the ontological structure defined in Figure 2 to model the change objects and their LULC in our study area. The knowledge in the ontology is generated automatically utilizing the semantic framework introduced in Section 3.3. In Figure 5, we demonstrate the visualized presentation of the class (Figure 5a), property (Figure 5b) and individual (Figure 5c,d) definitions of the ontology fragment. The ontology contains five classes: change object, satellite image, spatial dimension, temporal dimension and thematic dimension.

The "Change\_Object" class is defined to organize the geospatial objects that contain LULC change over time. "Village" in the context of this study is defined as its subclass. The "Satellite\_Image" class contains key information to describe image data used in the LULCC analysis. It contains three subclasses: (1) "Image\_Feature", which refers to various low-level features that an image has, such as color, shape or texture; (2) "Image\_Space" describes the spatial extent and resolution of a satellite image; and (3) "Meta\_information" includes the name for the image, the sensor used to acquire the image data and the timestamp at which an image was taken. Three other classes defined in this ontology are the spatial dimension (location), temporal dimension and thematic dimension of a change object.

Besides the class definition, we also define two types of properties: object property and data property. The object property is mainly used to connect a subject and object, which are both individuals of a certain class in a "subject-predicate-object" triple structure. As shown in Figure 5b, various topological relationships are defined as object properties to represent the relationship between change objects. When the value range of a property (also known as the predicate) is a native data type, such as numbers, strings or time, supported by the OWL standard, they are defined as a data property.

After defining classes and properties, the ontology is instantiated. Figure 5c,d demonstrates the triple definition for the QuickBird image (which is an individual of the "Satellite\_Image" class) and an example change object named "2010\_R1" segmented from the image. For the image, we define its name, geographical extent, sensor name, spatial resolution and the coordinate system used for data processing. For the change object "2010\_R1", its area, boundary geometry, existence time, identity and LULC type are encoded. Note that we do not hardcode topological relationships; instead, this knowledge is dynamically acquired through a semantic reasoning process based on known knowledge in the ontology. The reason for this is that the knowledge about topological relationships is always too huge to model manually, especially when the study area is large and has undergone dramatic land use change over time. For instance, if there are n objects in one image scene at one timestamp, the triples defining the topological relationships (total number: m) among these spatial objects could become as many as  $m*n^2$ . Defining them manually is therefore not a scalable approach [10]. To ensure a successful and dynamic reasoning process, every change object is encoded by its geometry information (see triples with property "hasBoundary\_Geometry" in Figure 5d). The integrated semantic

and spatial reasoning framework makes a spatial inference and derives the topological relationships among different change objects on the fly.



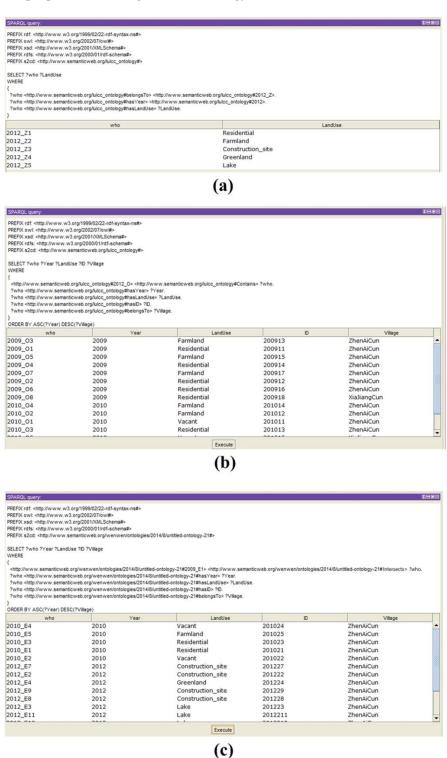
**Figure 5.** A visualized ontology segment for geospatial semantic modeling demonstration. (a) Defines all of the classes; (b) defines the data properties (such as the metadata information about an image, the spatial and temporal properties of a spatial object), as well as the object property (such as the topological relations among different objects); (c) a visual definition of triples for a QuickBird image from which the spatial objects are extracted; (d) a spatial object named "2010\_R1" and its various properties. For instance, the spatial reference information is defined in "hasCRS" (Coordinate Reference System).

# 4.3. Change Information Query Based on Semantic Reasoning

In order to test our geospatial semantic model, we designed a series of semantic queries based on the LULCC ontological model using Wantou data as a case study. The three queries discovered a spatial object's current LULC information, the historical LULCC trajectory and its LULCC trend(s). The query language is encoded in the semantic web query language SPARQL. Below, we pose three questions, each representing a query to demonstrate the land use reasoning supported by the reasoning tool.

Since there has been a very complex variation of land use in Wantou Community, we select three change objects for demonstration. First, we select ZhenAiCun (ID: 2012\_Z) in 2012 to show its current LULC information. Then, the historical LULCC trajectory for the geospatial object Xinhu Lake (ID: 2012\_O) before 2012 is shown. In 2012, the northern part of the lake 2012\_O1 belongs to ZhengAiCun, and the southern part 2012\_O2 belongs to XiaJiangCun. Finally, we show the trend of LULCC at geospatial object Miaoqianzhou (ID: 2009\_E1) from 2009 to 2012.

The query results, given in Figure 6, show that our semantic modeling method can support semantic geospatial queries for the change information of remote sensing imagery. In other words, the rules, restrictions, present condition, historical trajectory and tendencies of any change object or change object set can be found using a combination of different properties and relations (or classes, individuals and properties in Protégé) in our ontology.



**Figure 6.** Semantic query results using SPARQL. (a) The LULC composition of ZhenAiCun (2012\_Z) in 2012; (b) the historical trajectory of LULC of a spatial object with name "Xinhu Lake"; (c) the semantic query and result for the LULCC trend at Miaoqianzhou (2009\_E1) from 2009 to 2012.

The result of the SPARQL query in Figure 6a shows that the composition of LULC at ZhenAiCun (2012\_Z) in 2012 includes residential land (2012\_Z1), farmland (2012\_Z2), construction site (2012\_Z3), green land (2012\_Z4) and lake (2012\_Z5). Figure 7a shows its land use map. The result of the SPARQL query in Figure 6b illustrates the historical trajectory of LULC of Xinhu Lake (2012\_O) between 2009 and 2012. The changes over the period from 2010 to 2012 are illustrated using orange lines, and the changes from 2009 to 2010 are shown using pink lines in Figure 7. In 2012, Xinhu Lake (2012\_O) is partially within ZhenAiCun and XiaJiangCun. Actually, the part of Xinhu Lake within ZhenAiCun (2012\_O1) was once vacant land (2010\_O1), farmland (2010\_O2 and 2010\_O4) and residential land (2010\_O3) in 2010. The part within XiaJiangCun (2012\_O2) was actually developed from a vacant parcel (2010\_O5) in 2010. Tracing back to 2009, the vacant land (2010\_O1) in ZhenAiCun was residential in 2009. The farmland and residential area in the southern part remain mostly unchanged. While the southern part of the Xinhu Lake in XiaJiangCun in 2009 is completely residential, it became vacant (2010\_O1) in 2010. There are two major residential areas in both villages that turned to vacant land awaiting construction, indicating that there is substantial reconstruction between 2009 and 2010.

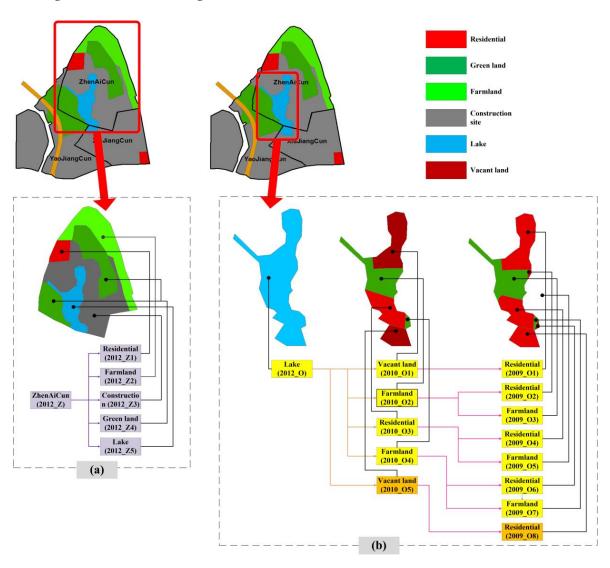
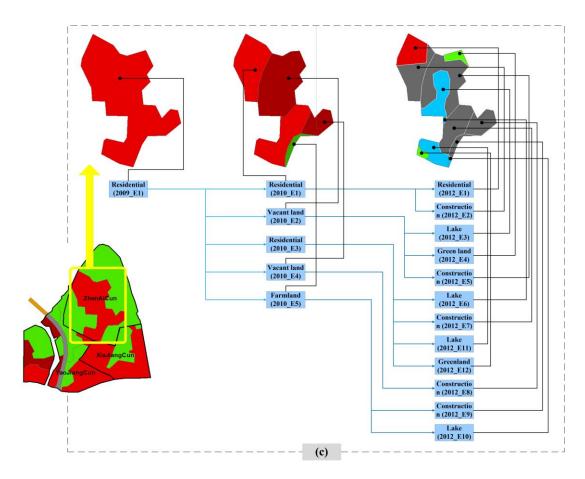


Figure 7. Cont.



**Figure 7.** Visual presentation of the semantic query results in Figure 6. (a) Land use of ZhenAiCun (2012\_Z) in 2012; (b) the historical trajectory of land use of Xinhu Lake (2012\_) from 2009 to 2012; and (c) the changed tendency of land use of Miaoqianzhou (2009\_E1) from 2009 to 2012.

The result of the SPARQL query in Figure 6c illustrates the LULCC trend at Miaoqianzhou's (2009\_E1) from 2009 to 2012. Changes from 2009 to 2010 are represented using light blue lines, and changes from 2010 to 2012 are represented using dark blue lines in Figure 7. In 2009, the entire Miaoqianzhou area is residential. In 2010, about half of Miaoqianzhou was torn down and became vacant land (2010\_E2 and 2010\_E4). There is a small portion in the southwest that turned into a farmland (2010\_E5). In 2012, a large part of Miaoqianzhou (2012\_E2, 2012\_E5, 2012\_E7, 2012\_E8 and 2012\_E9) was under construction. Parts of the vacant land 2010\_E2 in 2010 were developed into lake area (2012\_E6 and 2012\_E11) and green land (2012\_E12). It is apparent that the majority of construction was completed in 2012, and a new community had been established. The detailed relationships in these three query results are shown in Figure 7.

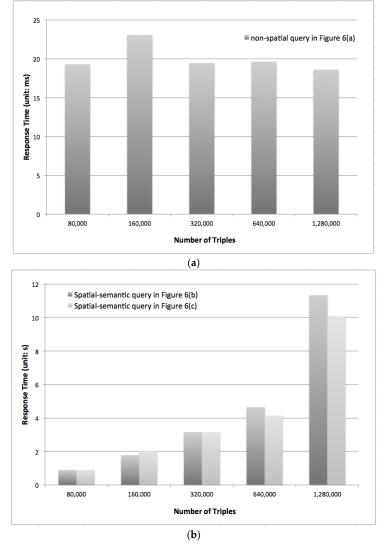
In summary, with the support of our semantic model and reasoning, it is simple and easy to obtain land use changes moving both backwards and forwards from a point in time.

#### 5. Experiments

To validate the scalability of the proposed software framework, we further conducted a series of experiments using simulated datasets, which include polygonal spatial objects increasing from 5000 to 80,000 at a factor of 2. In correspondence, after encoding the spatial, temporal and image properties for these objects, the triple sizes become 80,000, 160,000, 320,000, 640,000 and 1,280,000, respectively, for the simulated datasets. The same non-spatial semantic query (presented in Figure 6a) and the spatiotemporal and semantic queries (presented in Figure 6b,c) were tested using the big dataset. Our experiments were conducted on a Dell PowerEdge server with 8 G memory, one 6-core, 3.4-GHz CPU, 1-TB storage disk and an Ubuntu Server v14.04 operating system. The Jena bundled

in Apache software Apache-Jena-3.0.0 was used to implement the semantic platform, which enables web access.

Figure 8 shows the response time averaged from a hundred runs for each query to eliminate outlier results due to a network problem or a sudden system overuse. As seen, for the non-spatial semantic query about land use composition (the Figure 6a case), the system response time is very fast (below 25 ms) for all triple sets (see Figure 8a). The fact that the response time is not affected by the size of triples validates the scalability of the developed system to handle such semantic queries. This good result is attributed to the Apache Lucene full text indexing technique introduced in the Web Jena of our semantic platform. The results for spatial queries, in comparison, demonstrate different patterns. In the two spatial queries we tested, one (Figure 6b case) involves the spatial operator "contains", and the other (Figure 6c case) involves the spatial operator "intersects". It can be observed and concluded from Figure 8b that: (1) the response time of these two spatial queries are almost the same; (2) because spatial queries involves more computation, their response time is slower than non-spatial queries, which can normally be done instantly; (3) although it takes a longer time to obtain the results for complex spatial queries, the system still scales well; only a linear increase is observed: as the number of triples involved in the computation doubles, the response time gets doubled approximately. These experiments show the good scalability of the proposed system in answering semantically both spatial and non-spatial questions.



**Figure 8.** Response time of the semantic platform for both non-spatial and spatial semantic queries. (a) Response time for non-spatial semantic query; (b) response time for spatial-semantic queries.

#### 6. Conclusions

This paper introduces the development and implementation of a computational framework to support the modeling, semantic and spatial reasoning of change information with regard to space, time and topology. The center of the system is a reasoning engine capable of linking existing knowledge and semantic rules to infer dynamically new knowledge for spatial thinking. The knowledge base that the system relies on is encoded in an ontology that models various properties (i.e., image, spatial, temporal) of a single spatial object and the spatial relationships among different objects. Earlier studies of LULCC modeling lack a knowledge-based approach to connect the results of change detection with geoinformation in GIS or spatial databases. Our work integrates both aspects of change information into the proposed ontological model by transforming image, spatial, temporal and thematic semantics in the semantic LULCC model. In addition, this model is not limited to use in LU studies; it is generally applicable for other applications, such as forest management, climate changes and fire monitoring that require extensible use of geographical data. Another important contribution of this work is the proposal of the ways to integrate the machine-understandable ontology, reasoning rules, spatiotemporal data, as well as the inference engine seamlessly into an operational software framework to support on-the-fly query, reasoning and visualization of the change information.

There are several possible extensions of this research. First, we are implementing and integrating the proposed spatial ontology and reasoning framework into an operational cyberinfrastructure framework to support collaborative LULCC querying and decision-making. A prototype has been established [49]. In order to address the computational challenges in handling the management and reasoning of (billion-level) big LULCC data, we are extending the backbone reasoning model to develop a scalable semantic framework that can provide high performance support to spatial indexing, querying and reasoning. Finally, the ontological definition of classes and instances in this work is centered on discrete geospatial objects. This allows us to extend this ontological model to formally describe the geographical scenes that contain both objects and relationships between objects. We believe our work makes a major contribution in terms of providing both conceptual and practical solutions to the semantic modeling and reasoning of change data, and it will greatly benefit the broader Semantic Web and GIScience communities.

Acknowledgments: This research is supported in part by USGS G15AC00085 and NSF-BCS 1455349.

**Author Contributions:** Wenwen Li and Xiran Zhou conceived of the idea of the paper. Xiran Zhou and Sheng Wu performed the experiments. Wenwen Li and Xiran Zhou analyzed the data. Sheng Wu and Wenwen Li developed the semantic reasoning platform. Wenwen Li and Xiran Zhou wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- 1. Bambacus, M.; Yang, C.; Evans, J.; Li, Z.; Li, W.; Huang, Q. Sharing earth science information to support the Global Earth Observing System of Systems (GEOSS). In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 6–11 July 2008.
- 2. Yang, C.; Li, W.; Xie, J.; Zhou, B. Distributed geospatial information processing: Sharing distributed geospatial resources to support Digital Earth. *Int. J. Digit. Earth* **2008**, *1*, 259–278. [CrossRef]
- 3. Liu, B.; Wang, M.; Hong, R.C.; Zha, Z.J.; Hua, X.S. Joint learning of labels and distance metric. *IEEE Trans. Syst. Man Cybern Part B* **2010**, 40, 973–978. [CrossRef] [PubMed]
- 4. Li, W.; Yang, C.W.; Nebert, D.; Raskin, R.; Houser, P.; Wu, H.; Li, Z. Semantic-based web service discovery and chaining for building an Arctic spatial data infrastructure. *Comput. Geosci.* **2011**, *37*, 1752–1762. [CrossRef]
- 5. Albalawi, E.K.; Kumar, L. Using remote sensing technology to detect, model and map desertification: A review. *J. Food Agric. Environ.* **2013**, *11*, 791–797.
- 6. Klein, I.; Dietz, A.J.; Gessner, U.; Galayeva, A.; Myrzakhmetov, A.; Kuenzer, C. Evaluation of seasonal water body extents in Central Asia over the past 27 years derived from medium-resolution remote sensing data. *Int. J. Geogr. Inf.* **2014**, *26*, 335–349. [CrossRef]

- 7. Kennedy, R.E.; Townsend, P.A.; Gross, J.E.; Cohen, W.B.; Bolstad, P.; Wang, Y.Q.; Adams, P. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sens. Environ.* **2009**, *113*, 1382–1396. [CrossRef]
- 8. Hamandawana, H.; Eckardt, F.; Chanda, R. Linking archival and remotely sensed data for long-term environmental monitoring. *Int. J. Appl. Earth Obs. Geoinform.* **2005**, *7*, 284–298. [CrossRef]
- 9. Matchanov, M.; Teodoro, A.; Schroder, C. Criterion definition for the identification of physical-geographical boundaries of Khorezm oasis through remotely sensed data. *Environ. Monit. Assess.* **2016**, *188*, 1–14. [CrossRef] [PubMed]
- 10. Arenas, H.; Harbelot, B.; Cruz, C. LC3: A spatiotemporal data model to study qualified land cover changes. In *Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects*; Ahlqvist, O., Varanka, D., Fritz, S., Janowicz, K., Eds.; CRC Press: Boca Raton, FL, USA, 2015; pp. 211–242.
- 11. Shi, X. Where are the spatial relationships in the spatial ontologies? *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E459. [CrossRef] [PubMed]
- 12. Gribb, W.J.; Czerniak, R.J. Land use/land cover classification systems and their relationship to land planning. In *Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects*; CRC Press: Boca Raton, FL, USA, 2015; pp. 1–18.
- 13. Lu, D.; Mausel, P.; Brondizio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, 25, 2365–2407. [CrossRef]
- 14. Ahlqvist, O. Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: A study of 1992 and 2001 US National Land Cover Database changes. *Remote Sens. Environ.* **2008**, *112*, 1226–1241. [CrossRef]
- 15. Liu, Z.G.; Dezert, J.; Mercier, G.; Pan, Q. Dynamic evidential reasoning for change detection in remote sensing images. *IEEE Trans. Geosci. Remote* **2012**, *50*, 1955–1967. [CrossRef]
- Björk, A.; Skånes, H. The need for awareness of semantic plasticity in international harmonization of geographical information: seen from a nordic forest classification perspective. In *Land Use and Land Cover Semantics: Principles, Best Practices and Prospects*; CRC Press: Boca Raton, FL, USA, 2015; pp. 41–58.
- 17. Anderson, J.R. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*; US Government Printing Office: Washington, DC, USA, 1976.
- 18. European Environment Agency (EEA). CORINE Land Cover Technical Guide: Addendum 2000. Available online: http://www.eea.europa.eu/publications/tech40add (accessed on 10 February 2016).
- 19. American Planning Association. Land-Based Classification Standards. 2001. Available online: https://www.planning.org/lbcs/standards/pdf/InOneFile.pdf (accessed on 10 February 2016).
- 20. Di Gregorio, A. Land Cover Classification System: Classification Concepts and User Manual. Available online: http://www.fao.org/docrep/008/y7220e/y7220e00.HTM (accessed on 10 February 2016).
- 21. Harrison, A.R. *National Land Use Database: Land Use and Land Cover Classification*; O. Office of the Deputy Prime Minister, Ed.; Queen's Printer and Controller of Her Majesty's Stationery Office: London, UK, 2006.
- 22. Jansen, L.J. Parameterized approaches to the categorization of land use and land cover. In *Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects;* CRC Press: Boca Raton, FL, USA, 2015; pp. 59–84.
- 23. Annoni, A. INSPIRE: Infrastructure of spatial information in Europe. In Proceedings of the 9th EC GIS&GI A, Coruna, Espana, 25–27 June 2003.
- 24. Rips, L.J.; Shoben, E.J.; Smith, E.E. Semantic distance and the verification of semantic relations. *J. Verbal Learn. Verbal Behav.* **1973**, 12, 1–20. [CrossRef]
- 25. Tversky, A. Features of similarity. *Psychol. Rev.* 1977, 84, 327–352. [CrossRef]
- 26. Li, W.; Raskin, R.; Goodchild, M.F. Semantic similarity measurement based on knowledge mining: An artificial neural net approach. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1415–1435. [CrossRef]
- 27. Sunna, W.; Cruz, I.F. Structure-based methods to enhance geospatial ontology alignment. In *GeoSpatial Semantics*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 82–97.
- 28. Goldstone, R.L. Similarity, interactive activation, and mapping. *J. Exp. Psychol. Learn. Mem. Cogn.* **1994**, 20, 3–28. [CrossRef]
- 29. Lin, D. An information-theoretic definition of similarity. ICML 1998, 98, 296–304.
- 30. Comber, A.; Fisher, P.; Wadsworth, R. Text mining analysis of land cover semantic overlap. In *Land Use* and *Land Cover Semantics: Principles, Best Practices, and Prospects*; CRC Press: Boca Raton, FL, USA, 2015; pp. 197–204.

- 31. Hornsby, K.; Egenhofer, M.J. Identity-based change: A foundation for spatio-temporal knowledge representation. *Int. J. Geogr. Inf. Sci.* **2000**, *14*, 207–224. [CrossRef]
- 32. Worboys, M.F. Modelling changes and events in dynamic spatial systems with reference to socio-economic units. In Proceedings of the ESF GISDATA Conference on Modelling Change in Socio-Economic Units, Napthlion, Greece, 15–19 March 1998.
- 33. Claramunt, C.; Jiang, B. An integrated representation of spatial and temporal relationships between evolving regions. *J. Geogr. Syst.* **2001**, *3*, 411–428. [CrossRef]
- 34. Jiang, J.; Worboys, M. Event-based topology for dynamic planar areal objects. *Int. J. Geogr. Inf. Sci.* **2009**, 23, 33–60. [CrossRef]
- 35. Varanka, D.E.; Usery, E.L. An applied ontology for semantics associated with surface water features. In *Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects*; CRC Press: Boca Raton, FL, USA, 2015; pp. 146–168.
- 36. Prasad, A.R.D.; Guha, N. Concept naming vs. concept categorisation: A faceted approach to semantic annotation. *Online Inf. Rev.* **2008**, *32*, 500–510. [CrossRef]
- 37. Batsakis, S.; Petrakis, E.G. Representing temporal knowledge in the semantic web: The extended 4d fluents approach. In *Combinations of Intelligent Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 55–69.
- 38. Klyne, G.; Carroll, J.J. Resource Description Framework (RDF): Concepts and Abstract Syntax; W3C: Cambridge, MA, USA, 2006.
- 39. McGuinness, D.L.; Van Harmelen, F. OWL web ontology language overview. W3C Recomm. 2004, 10, 1–22.
- 40. Gutierrez, C.; Hurtado, C.; Vaisman, A. Introducing time into RDF. *EEE Trans. Knowl. Data Eng.* **2007**, 19, 207–218. [CrossRef]
- 41. Frasincar, F.; Milea, V.; Kaymak, U. tOWL: Integrating time in OWL; Springer: Berlin/Heidelberg, Germany, 2010.
- 42. Krieger, H.U. A General Methodology for Equipping Ontologies with Time. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.168.1549&rep=rep1&type=pdf (accessed on 23 June 2016).
- 43. Bereta, K.; Smeros, P.; Koubarakis, M. Representing and querying the valid time of triples for Linked Geospatial data. In Proceedings of the 10th Extended Semantic Web Conference (ESWC 13), Monpellier, France, 26–30 May 2013.
- 44. Huang, Y.; Deng, G. Research on representation of geographic spatio-temporal information and spatio-temporal reasoning rules based on geo-ontology and SWRL. In Proceedings of the International Conference on Environmental Science and Information Application Technology, Wuhan, China, 4–5 July 2009.
- 45. Sotirios, B. SOWL: A Framework for Handling Spatio-Temporal Information in OWL. Ph.D. Thesis, Technical University of Crete, Chania, Greece, 2011.
- 46. Bledsoe, B.P.; Brown, M.C.; Raff, D.A. GeoTools: A toolkit for fluvial system analysis. *J. Am. Water Resour. Assoc.* **2007**, 43, 757–772. [CrossRef]
- 47. McBride, B. Jena: A semantic web toolkit. IEEE Int. Comput. 2002, 6, 55–59. [CrossRef]
- 48. Prud'Hommeaux, E.; Seaborne, A. SPARQL Query Language for RDF. Available online: http://www.w3. org/TR/rdf-sparql-query/ (accessed on 23 June 2016).
- 49. PolarGlobe. Available online: http://polar.geodacenter.org/lulcc (accessed on 26 September 2016).



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).