# Principles of 5D modeling, full integration of 3D space, time and scale

Peter van Oosterom and Jantien Stoter
Delft University of Technology, OTB, GIS technology
Jaffalaan 9, 2628 BX Delft, The Netherlands
{p.j.m.vanoosterom|j.e.stoter}@tudelft.nl

## Abstract

This paper proposes an approach for data modelling in five dimensions. Apart from three dimensions for geometrical representation and a fourth dimension for time, we identify scale as fifth dimensional characteristic. Considering scale as an extra dimension of geographic information, fully integrated with the other dimensions, is new. Through a formal definition of geographic data in a conceptual 5D continuum, the data can be handled by one integrated approach assuring consistency across scale and time dimensions. Because the approach is new and challenging, we choose to step-wise studying several combinations of the five dimensions, ultimately resulting in the optimal 5D model. We also propose to apply mathematical theories on multidimensional modelling to well established principles of multidimensional modelling in the geo-information domain. The result is a conceptual full partition of the 3Dspace+time+scale space (i.e. no overlaps, no gaps) to be realised in a 5D data model implemented in a Database Management System.

## 1 Motivation

The role of geographic information in our society has changed tremendously the past decades. From being collected and used ad hoc for maps and other specific applications, it has now become part of the Geo Information Infrastructure (GII). Ultimately GIIs will serve complete information flows in the web from observing/monitoring via processing/analyzing/planning to communicating and controlling actions. After many isolated initiatives, our society is now heading towards a sustainable GII in which geo-data is shared and re-used by many and highly different users and applications through machine-based linking of large amounts of distributed geographic data and information.
Formal definitions of geo-information are required to enable understanding and satisfying the requests of people and (increasingly) machines to use appropriate geo-information available within the web. This paper focuses on formalising the data models and structures that capture geo-information. Data structures for geo-information bring specific challenges as traditional DataBase Management System (DBMS) implementations for non-dimensional information are not capable of handling the different dimensions of geo-data sufficiently. In our approach we distinguish five dimensional concepts of geo-data.
Apart from their location and 0D to 3D geometrical and topological characteristics, geo-data has further temporal (when was a moving object at a specific location; when was an object valid in the database?) and scale components that were often implicitly taken into account when the data was collected. These different dimensional aspects highly correspond, e.g. a (possibly geometric) change may be only relevant for the highest scale of an object or understanding the route directions for a long car trip requires overview, but at specific locations (e.g. to rest or to stay overnight) consistent information at a higher scale, with also temporal information (i.e. weather conditions at a certain moment) may be needed. Although (multi-)scale is a well-known concept in the geo-information technology domain, regarding it as an extra dimension of geo-data, integrated with the other dimensions, is new.

Despite the interdependencies, until now different dimensions of geo-data have been studied in separate initiatives, with sometimes limited support for the other dimensions. Although these past studies have gained important knowledge on how to handle the individual dimensions 2D/3D space, time and scale, no modelling approach exists that truly integrates all dimensional concepts of geo-data at a fundamental level. This is our motivation to start a new research on a conceptual full partition of 3Dspace+time+scale (i.e. no overlaps, no gaps) realised in a true 5D generic model. This paper elaborates on the research methodology that we propose to accomplish the true 5D model. In contrast to a separate handling of spatial, temporal and scale dimensions, a true 5D approach provides a sustainable and solid foundation for the GII for three main reasons:

- The deep integration of all dimensional concepts accomplishes a highly formal definition of geo-data (with 5D data types and 5D topological primitives) as the relationships between space, time and scale aspects of geo-data are fully addressed and no special cases need to be treated in another way anymore. Every case will be a specialisation of the model.
- The model enforces consistency crossing dimensional borders which improves the quality of geo-data.
- Optimal efficient 5D searching and maintenance can only be realised if a 5D data type and index/clustering is used, otherwise the DBMS query plan has to select first on space, then on time and then on scale (or in another order). An example of a 5D search that appears in space, time, and scale context is the integration of a database with physical plans at different moments in history and a database with historical information on buildings to check which buildings (extensions) conflict with which status of the physical plan. Another example of a 5D search is comparing the cadastral database that registers the legal status of networks based on the physical extent of the network at the moment of registration with the physical registration maintained by the network company in which changes as extensions, deletion and movements of parts of the network are recorded as well.

In our approach the multidimensional integration is studied at two levels. At first a conceptual 5D data model will be designed on which all other geo-data models can be founded. Secondly, as the foundation of the GII consists of geo-DBMSs maintaining geographic information, the methodology proposes to study how DBMS functionality can be extended up to 5D as implementation of the conceptual model.
We do recognise the high ambitions to realise a true 5D data model. However our aim is to lay down a foundation for multidimensional data modelling by defining a theory validated through prototype implementations, which can be further developed in the future. In addition, in a step-wise approach we will apply mathematical theories on multidimensional modelling to established principles in 2D/3D, time and multi-scale modelling. By studying several combinations of the different dimensions, we will accomplish the optimal method for including multidimensional concepts and notions in geo-data modelling. This paper is a condensed version of our earlier publication (van Oosterom and Stoter, 2010).
To explain our proposed methodology in which we combine multidimensional principles established in both the geo-domain and mathematical theories, we first elaborate on the multidimensional modelling concepts in the geo-information domain (section 2), while in section 3 we will explain the potential mathematical theories on multidimensional modelling that we will explore. Section 4 explains our proposed methodology in more detail and the paper ends with discussion in section 5.

# 2 Ingredients of nD modeling: space, time and scale.

Several previous researches have studied multidimensional modelling of geo-information. A first related topic is nD storage and mining (Gray et al., 1997, Casali et al., 2003) which aims at aggregating information on multiple thematic (non-spatial) attributes to perform efficient database querying subsequently. For example 5D data would be the result of combining object-id, weight, colour, price, and date attributes. Although this 5D data focuses on thematic attributes in data mining and not on dimensional concepts in geo-data modelling, the similarity is that it also considers multiple aspects of

data in an integrated manner. Note that this is nD point data and does not address the higher dimensional geometric primitives (lines, areas, volumes) and the true integration with time and scale.
Also the particular domain of modelling is relevant (Hamilton et al., 2005, Aouad et al., 2005). For example, nD modelling extends BIMs (Building Information Models) with additional thematic characteristics to serve each stage of the lifecycle of a building facility through one common information model. However nD modelling focuses mainly on integrating multiple thematic, and not multidimensional, concepts. As the multidimensional data model that we propose offers a framework to structure any geo-information, the thematic approach of nD modelling can be well served by progresses in multidimensional data modelling, for example 4D BIMs that include the time dimension.

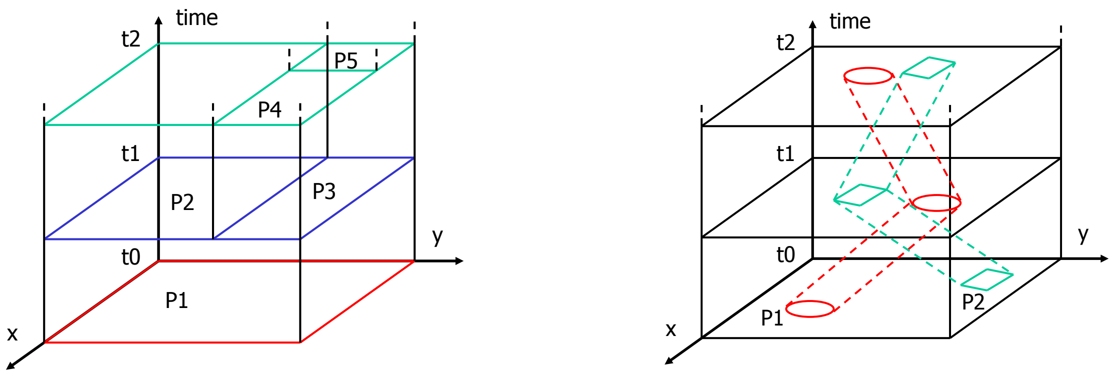


Fig. 1: Time as third dimension: division of parcels (left) and moving objects (right) (van Oosterom et al., 2006)

Temporal aspects of geo-data is fundamental for recording or monitoring changes, for describing processes, and for documenting future plans. For example monitoring the status change of a set of related features (Figure 1, left) or monitoring changes of moving objects (Figure 1, right). Many Spatio-Temporal (ST) data models have been designed to model changes of geo-information (Hornsby and Egenhofer, 2000, Peuquet, 2002, Raper and Livingstone, 2001). The semantics of the time dimension included in these models vary from model to model. The deep integration of time with space and scale concepts will fully handle changes upon position, attributes and/or extent of the objects in the unified space-time-scale continuum. Some aspects require specific attention for the time dimension in this deep integration. At first all possible changes of geo-information at varying scales should be well represented, i.e. change in geometry OR topology OR attributes, or no change at all. In addition the integrated representation of time should not only support changes at discrete moments, as currently supported by most of the ST models via timestamps and versioning, but also continuous temporal changes to describe the movement or change of objects independently from their object identification. Also the integrated space-time-scale approach requires specific attention for topological relationships between (continuously) evolving geographic objects. More researchers have identified the need for a generic spatio-temporal data model. A theory on a unified spatial-temporal data model was proposed in (Worboys, 1994) and a generic spatio-temporal data type in a relational DBMS was suggested in (Jin et al., 2005, van Oosterom et al., 2002).
Modelling different scales of geo-data is related to the “coarse-to-fine” hierarchical structure of how we perceive, model and understand our environment. In some applications less detailed, but simpler data works better, especially when there is a need for an overview. In other cases very detailed data is required. Two basic options exist for maintaining data sets of the same real world at different scales. First option is to separately maintain different databases at predefined scale-steps. This option is practiced by

many National Mapping Agencies that produce maps at different scales. Second option is to maintain only the most detailed data and to automatically generalise small scale data from it on the fly, eventually supported by pre-storing the results of costly geometric computations in multi-representation (as in the first option). To provide and reuse multi-scale data within the GII, consistency between data at different scales is fundamental, i.e. the availability of data at different scales free from contradictions enabling smoothly zooming in and out. This is supported by multi-representation data models that formally define different scale states of the data. Many researchers have studied multi-representation data models since it was introduced in (NCGIA, 1989, Buttenfield and Delotto, 1989). Examples are MRMS (Friis-Christensen and Jensen, 2003), MADS (Parent et al., 2006), Perceptory (Bédard, 2004), modelling multiple geometries (Jones et al., 1996), modelling scale transitions between pairs of objects (Devogele et al., 1996) and modelling links between instances (Kilpelainen, 1997). While these previous initiatives mainly aimed at controlling the redundancy of multi-representations and multi-scale data, our study will specifically aim at reducing redundancy to improve efficiency and to better assure consistency between different scales: vario-scale structures (van Oosterom, 2006, van Oosterom and Meijers, 2011, Meijers, 2011).

# 3 Mathematical Theories on Multidimensional Modelling

To realise a 5D geo-data modelling approach by which the treatment of (up to 3D) space, time and scale is optimally integrated, we will study existing mathematical theories on multidimensional descriptions and apply them to the well defined frameworks for 3D, time and scale modelling in the geo-information domain. Examples of established mathematical theories on multidimensional modelling are:

- Topological polyhedra where multidimensional objects are built from their lower primitives, i.e. a 3D volume object consists of 2D faces that consist of 1D edges that consist of 0D nodes (Croom, 2002, Cromwell, 1999).
- Regular polytopes, which is based on a division of space by hyperplanes, e.g. a 3D volume object is described by 2D planes (Coxeter, 1973, Thompson, 2007).
- Simplicial Homology based n-simplexes, which are the building blocks for the Triangular Irregular Network (TIN in 2D) and Tetrahedral Network (TEN in 3D) and their higher dimensional equivalents (Giblin, 1981, Penninga, 2008); see Figure 2.

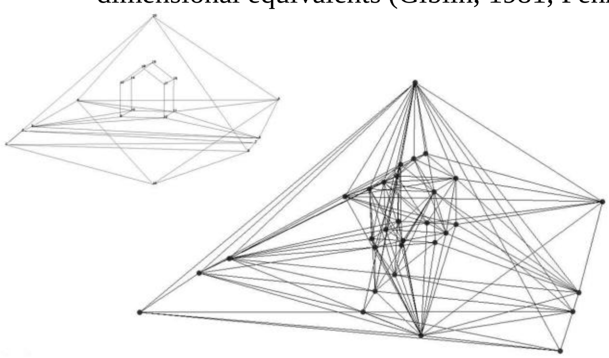

Fig. 2: Simple scene and the TEN [20].

The first theory is advantageous for multidimensional geo-data modelling because it aligns to the boundary representation of 3D volume objects of OGC (OGC, 1999). However since it lacks of a well defined fundament, validity of objects has to be fully handled by additional functionalities. The advantage of the second theory is that the formalisation of multidimensional concepts is straightforward because the primitives that build an object are described with equations of the hyperplanes (and are valid in any dimension). Finally, the third theory is advantageous for geo-data modelling because of the n-simplex (e.g. triangle, tetrahedron) based approach. Triangles contain specific characteristics, such as convexity, that make it easy to enforce validity of objects that consist of the lower dimensional primitives (and again this theory is valid in any dimension).

# 4 Research methodology for 5D Data Modelling

Because of the unexplored domain of deeply integrated 5D information modelling, much knowledge need to be gained on the optimal 5D approach. To do so, in our methodology we propose to first apply mathematical theories on multidimensional modelling to principles established in 2D/3D, spatio-temporal and multi-scale data models and to gradually extend the results with extra dimensions in three iterations (A, B and C), see Figure 3. This will lead to three alternative 3D models in the first iteration (3Dspace, 2Dspace+time, 2Dspace+scale) and three alternative 4D models in the second iteration (3Dspace+time, 3Dspace+scale, 2Dspace+time+scale), finally leading to the best 5D data model in which lower dimensional objects are supported as well. The intermediate trajectory is important to optimally prepare the separate approaches for an integrated 5D data modelling approach and to gain fundamental knowledge on how to best address the different dimensions in the integration, both at conceptual model level and on database technology level. The steps applied in every iteration are: 1. Conceptual modelling, 2. Implementation (with test data), and 3. Testing and validation (with real world scenarios).

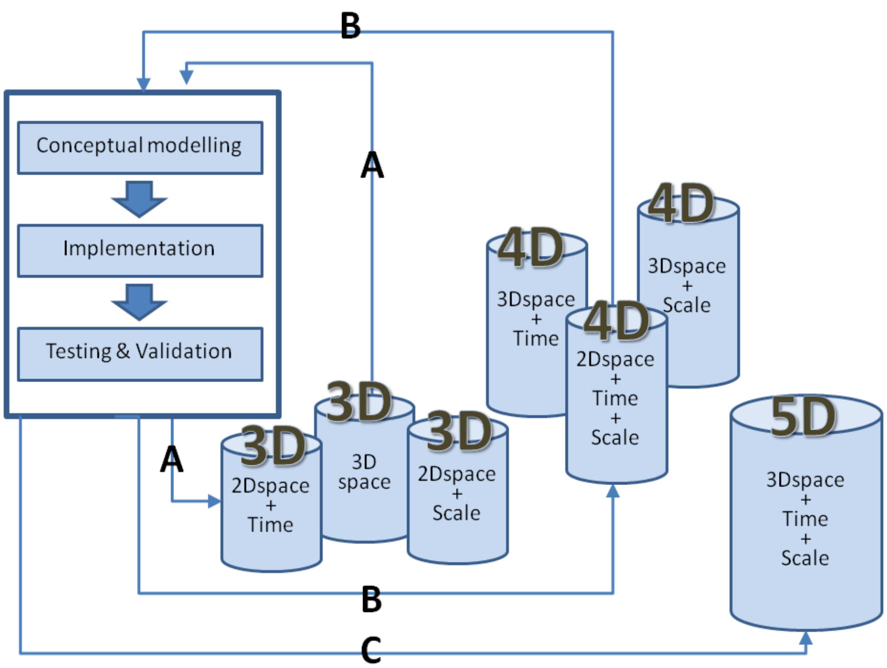


Fig. 3: Workflow of research methodology

Since iteration A starts from established principles, the resulting 3D models should be well feasible and in reach, also because they can be built on the partial models that are already operational, e.g. 3D in commercial systems (Bentley Systems, Oracle, ESRI), spatio-temporal databases and vario-scale data structures. Therefore the main aim of this iteration is gaining insight into how integrated spatial and time (or scale) dimensions behave when applying the multidimensional mathematical theories (simplicial homology, regular polytopes). The combined dimensions of iteration B belong to hardly explored types of models and implementations, but we expect them to be feasible (in part). In addition they will provide insight in even more complicated integrated 5D modelling, which is still required as these 4D models focus on a selection of multidimensional concepts only. Again these explorations will provide more insight in the behaviours of integrated dimensions. Iteration C will use the knowledge from iterations A and B to generate concepts for the 5D data model deeply integrating space, time and scale concepts implemented with a 5D data type, 5D topological structures and primitives as well as 5D clustering and indexing.

For validating the (intermediate) research results we will establish application tests with large datasets containing 2Dspace and 3Dspace geo-information at several scales also containing time information. For these tests we can make use of the spatio-temporal cadastral database and of large, mid- and small-scale topographic datasets of the Netherlands' Kadaster. In addition various large 3D datasets are available, such as the Actual Height Model of The Netherlands, 3D datasets of municipalities (Amsterdam, Rotterdam and Tilburg) and a 3D detailed topographic dataset of Rijkswaterstaat.

# 5 Conclusion

This paper proposes a research methodology for 5D data modelling that fully integrates 2D/3D space, time and scale aspects of geo/information. The methodology combines established principles on 2D/3D space, time and scale modelling in the geo-information domain with mathematical theories on multidimensional modelling. Although 3D, time and scale aspects have been studied in separate research domains, studying the deep integration of time and scale concepts in the traditional 2D/3D models to replace their separate time and scale treatment with a full partition of 3Dspace+time+scale is new.

Several stakeholders will benefit from this research. A first group of stakeholders that will benefit are providers of geo-information for which the multidimensional data types provide important advantages with respect to efficiency and consistency compared to the current separate treatment of space, time and scale. A common characteristic of these providers is that they are responsible for maintaining and providing large amounts of geo-information at different scales for which it is increasingly important to keep history track record. A second group of stakeholders that will benefit from the intermediate and final results are vendors of geo-ICT systems that can implement the multidimensional data types as realised in prototypes. A final group of stakeholders, and perhaps in the long term the most important group, that will benefit from this research are end-users of geo-information who will be served by improved and new 5D aware applications and services.

In the long term results on 5D data modelling in the geo-information technology domain are important for standards on geo-information which are established and developed by ISO TC 211 and the Open Geospatial Consortium. The existence and relative importance of the classes in the various applications could also be considered in a more integrated manner as the sixth dimension: the semantic-dimension. For, the time being this is considered out of our research scope, because of the less universal nature of this/these additional 'dimension(s)' (comparable to the nD dimensional point data modeling as applied in data mining).

# Acknowledgements

This research is supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO) and partly funded by the Ministry of Economic Affairs, Agriculture and Innovation (Project codes: 11300 and 11185).

# References

Aouad, G., A. Lee, S. Wu (2005). Special issue on `From 3D to nD modelling´. Journal of Information Technology in Construction.

Bédard, Y, S. Larrivée S, M-J Proulx, M. Nadeau (2004). Modelling geospatial databases with plug-ins for visual languages: a pragmatic approach and the impacts of 16 years of research and experimentations on Perceptory. Berlin : Springer, 2004. LNCS 3289.

Buttenfield, B.P. and J.S. Delotto (1989). Multiple representations. Scientific Report for the Specialist Meeting. National Center for Geographic Information and Analysis (NCGIA), 1989. p. 87. Technical paper 89–3.

Casali, A., R. Cicchetti, and L. Lakhal (2003). The 3rd SIAM International Conference on Data Mining. Cube Lattices: a Framework for Multidimensional Data Mining. pp. 3004-3008.

Coxeter, H. S. M. (1973). Regular Polytopes. Dover Publications, 1973. p. 321. ISBN: 978-0486614809.

Cromwell, P.R. (1999). Polyhedra. Cambridge University Press, 1999. p. 460. ISBN: 0-521-66405-5.

Croom, F.H. (2002). Principles of Topology. 2. Uitgever Cengage Learning , 2002. p. 312. ISBN 9812432884.

Devogele T., J. Trevisan and L. Raynal (1996). Building a multi-scale database with scale transition relationships. International Symposium on Spatial Data Handling. pp. pp 337–351.

Friis-Christensen, C.S. and A. Jensen (2003). Object-relational management of multiply represented geographic entities. Cambridge, MA, USA, July 9–11, 2003. Proceedings of the Fifteenth International Conference on Scientific and Statistical Database Management.

Giblin, P.J. (1981). Graphs, Surfaces and Homology: An Introduction to Algebraic Topology. Second edition. New York : Chapman and Hall, 1981. ISBN-13: 978-0470989944.

Gray, J., S. Chaudhuri, A. Bosworth, A. Layman, D. Reichart, M. Venkatrao (1997). Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals. Kluwer Academic Publishers, 1997, Data Mining and Knowledge Discovery ) °c 199, Vol. 1, pp. 29-53. ISBN:1-58113-737-0.

Hamilton, A., H. Wang, A.M. Tanyer, Y. Arayici, X. Zhang X and Y. Song (2005). Urban information model for city planning, ITcon Vol. 10. Special Issue From 3D to nD modelling, 2005, ITcon, Vol. Vol 10, pp. pg. 55-67. http://www.itcon.org/2005/6.

Hornsby, K., and Egenhofer, M.J. (2000). Identity-based change: a foundation for spatio-temporal knowledge representation. International Journal of Geographical Information Science, Vol. 14, pp. 207–224.

Jin, P., L. Yue and Y. Gong (2005). Research on a Unified Spatiotemporal Data Model. Beijing, China: ISPRS Press, International Symposium on Spatial-temporal Modeling, Spatial Reasoning, Analysis, Data Mining and Data Fusion

Jones C.B., D.B. Kidner, L.Q. Luo, G.L. Bundy, J.M. Ware (1996). Database design for a multi-scale spatial information system. International Journal Geographic Information Science, Vol. 10, pp. 901-920.

Kilpelainen, T. (1997). Multiple representation and generalisation of geo-databases for topographic maps. PhD thesis. Finnish Geodetic Institute, 1997.

Meijers, M. (2011). Variable-scale Geo-information. PhD thesis Delft University of Technology, December 2011, 235 p. Published by Netherlands Geodetic Commission, Publications on Geodesy 77, Delft, 2011

NCGIA. National Center for Geographic Information and Analysis (1989). The research plan of the National Center for Geographic Information and Analysis. International Journal Geographical Information Systems, Vol. 3, pp. 117–136.

OGC. (1999). OpenGIS Simple Features Specification For SQL, Revision 1.1. [Online] 1999. http://www.opengeospatial.org/standards/sfs.

van Oosterom, P.J.M. (2006). Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest. Cartography and Geographic Information Science, Vol. 32, pp. 331-346.

van Oosterom, P.J.M., B. Maessen, and C.W. Quak (2002). Generic query tool for spatio-temporal data. International Journal of Geographical Information Science, Vol. 16, pp. 713–748.

van Oosterom, P. and M. Meijers (2011). Towards a true vario-scale structure supporting smooth-zoom. In: Proceedings of the 14th Workshop of the ICA Commission on Generalisation and Multiple Representation & the ISPRS Commission II/2 Working Group on Multiscale Representation of Spatial Data, 2011, Paris, 19 p.

van Oosterom, P.J.M. and H. Ploeger, J. Stoter, R. Thompson and C. Lemmen (2006). Aspects of a 4D Cadastre: A First Exploration. Munich, Germany, 2006. XXIII International FIG congress.

van Oosterom, P. and J. Stoter (2010). 5D Data Modelling: Full Integration of 2D/3D Space, Time and Scale Dimensions. In: S.I. Fabrikant, T. Reichenbacher, M. van Kreveld and M. Schlieder (Eds.); Proceedings of the Sixth International Conference GIScience 2010, Springer, pp. 311-324.

Parent, C, S. Spaccapietra and E. Zimányi (2006). Conceptual modelling for traditional and spatio-temporal applications. The MADS approach. Springer, 2006. ISBN: 3–540–30153–4.

Penninga, F. (2008). 3D Topography A Simplicial Complex-based Solution in a Spatial DBMS, PhD thesis. TU Delft. The Netherlands: Publications on Geodesy 66, 2008. p. 204. ISBN: 978 90 6132 304 4.

Peuquet, D.J. (2002). Representations of Space and Time. New York: Guilford, 2002. p. 394.

Raper, J.F. and D.E. Livingstone (2001). Let's get real: spatio-temporal identity and geographic entities. Transactions of the Institute of British Geographers, Vol. 26, pp. 237-42.

Thompson, R.J. (2007). Towards a Rigorous Logic for Spatial Data Representation. PhD thesis. Delft University of Technology. Netherlands Geodetic Commission, 2007. p. 333.

Worboys, M.F. (1994). A unified model for spatial and temporal information: Spatial data: applications, concepts, techniques. Oxford University Press, 1994, Computer journal, Vol. 37, pp. 26-34.