Modelling Uncertain Geodata for the Literary Atlas of Europe

Anne-Kathrin Reuschel, Barbara Piatti and Lorenz Hurni

Abstract Mapping literary spaces is an interdisciplinary challenge for both literary theory and cartographic realisation. The inherent properties of textual space compared to real-world geospace are introduced and discussed in relation to build a data model for the on-going project, 'Literary Atlas of Europe'. Spatial descriptions in fiction are often vague, transformed or hardly locatable within the real-world geospace which serves as basis for the analysis and visualisation. To reflect the fictional world, additional attributes and composed geometries are required. To solve the problem, an attributed spatial data model was developed to meet the requirements for a comparative, flexible study of literary spaces with numerous thematical questions. Using this model, spatial data of fictional texts were classified into settings, zones of actions, projected spaces, routes and marker. To emphasise the complexity of the model, the internal structure and diversity are illustrated with the examples of the spatial objects setting and route.

Keywords Uncertainty · Data acquisition · Data modelling · Literary geography · Spatial humanities

1 Introduction

Exploring space and place with the help of digital maps and Geographic Information Systems is a new trend that can be found more and more within the humanities. As formulated for scholars in humanities in an article recently

A.-K. Reuschel (✉) · B. Piatti · L. Hurni
Institute of Cartography and Geoinformation,
Swiss Federal Institute of Technology,
Zurich, Switzerland
e-mail: reuschel@karto.baug.ethz.ch

K. Kriz et al. (eds.), *Understanding Different Geographies*,
Lecture Notes in Geoinformation and Cartography,
DOI: 10.1007/978-3-642-29770-0_11, © Springer-Verlag Berlin Heidelberg 2013

published by Cohen (2011) "visualising data helps you to analyze it. The eye is a very good sorter of patterns". Cohen also comes to the conclusion that "humanities had become too abstract and neglected physical space. The value of what scholars are calling "*the spatial turn*" allows you to ask new questions: Why is it that something developed here and not somewhere else, what is it about the context of this place?" With similar ideas in mind, the interdisciplinary project of the 'Literary Atlas of Europe' aims to investigate fictionalised landscapes and cities. In doing so, scholars are interested in the development of the fictionalisation process of a specific region or in finding patterns while comparing the relation of fictionalised places to historical-political, natural (or other) events. To enable statistical investigations and comparable analyses and to consequently gain significant results of a region, a well-founded and if possible complete set of data about the region is required. To develop the methods and exploit the potential of such an interdisciplinary topic, a collection of texts was compiled in which each text is spatially linked to one of three model regions: an alpine landscape (Lake Lucerne/Gotthard in Switzerland), a coastal border area (Northern Frisia in Germany) and an urban space (Prague, Czech Republic).

To investigate the spatial characteristics of fictional texts, every single spatial occurrence needs to be extracted from the text. Within the frame of our project, spatial data of fictional texts are classified into five main categories: settings, zones of actions, projected spaces, routes and marker, as suggested by Piatti (2008). By applying these categories, the entire spatial structure of a literary text can be captured, while the function of the particular spatial entity is indicated at the same time. In the development process of the 'Literary Atlas of Europe' the question of how to collect and organise these data became a key issue.

These spatial data are descriptions by authors within a fiction. However, it is not absolutely necessary to have a spatial frame in each story, as Piatti noticed in her studies: "Actually, the poetic imagination doesn't need to stick to the physical space necessarily, neither to its geographical or topographical principles—but still, this is exactly what is done very often" (Piatti 2008—translated Reuschel). Of course, spatial data in fictions are transcribed in poetical words rather than geographical definitions like coordinate systems and scientifically defined variables.

Unlike spatial data models that attempt to measure, define and delimit boundaries of spatial objects as precisely as possible, based on data and facts, we are moving between objectivity and vague, subjective interpretation of indeterminate locations and their associated attributes. Typical descriptions of spatial data that can be found within fictions are, for instance: "in front of the gates of the city", "a villa on the first slopes of the steep hillside", "a house close to the Karl's bridge" or "a store in the main street". A further complication is that most geographical entities are not precisely delimited as Fisher (1999) notes in his research about *Models of Uncertainty in Spatial Data*. He argues that "well-defined geographical objects are essentially created by human beings to order the world they occupy. Most, if not all, […] geographical phenomena are […] poorly defined to some extent".

In the context of geographical phenomena the terms *vague* and *rough location* appear in various scientific publications and are substantiated with examples, that

could easily be found in the same way as in literary fictions. The authors use examples like the spatial concepts of "the Alps", "the Ruhr" and the "North Sea", more general concepts like "valleys", "dunes" and "forests" and vague transcriptions like "high mountains", "near the marsh" and "the rough end of the town" [compare Erwig and Schneider (1997), Bennett (2001), Bittner and Stell (2002) and Ahlqvist (2009)].

Montello et al. (2003) noticed in their paper called *Where's downtown*?, that "unlike formal languages, natural languages […] typically refer to categories that do not have precise references and are not delimited by sharp semantic boundaries". More geographical-specific terms of the notion *"natural languages"* were shaped and defined by Egenhofer and Mark (1995); (Waters and Evans 2003). The latter used the notion *"vernacular geography"* whereas Egenhofer shaped it to the term *"naive geography",* and realised: "many spatial inferences may appear trivial to us, (but) they are extremely difficult to formalize so that they could be implemented on a computer system".

Within the disciplines of geographic information sciences (GIScience), the term and concept of "uncertainty" is not uniformly defined and may draw on several different concepts such as error, accuracy, precision, completeness, confidence, and lineage. For a comprehensive discussion of the definition of uncertainty see e.g. MacEachren et al. (2005), Pang (2001), Thomson et al. (2005), Griethe and Schuhmann (2006) and Drecki (2007).

To classify uncertainty within literary geography, five uncertainty concepts are identified and listed here with examples from fictional analysis [concept definitions are adapted from Thomson et al. (2005), Waters and Evens (2003) and Fisher (1999)]:

Ambiguity	where geometry data are tied to linguistic expressions (e.g. "close" to a city, one of the neighbouring islands)
Averaging	where discrete geometry data are actually an average of time or scale-varying geographical descriptions (settings combined to form a zone due to the given scale e.g. various places in a town combined to form the town as a zone)
Continuousness	when geometry data are difficult to define because the measurements of an entity produce a gradient (e.g. the start of a setting in general, the start of a mountain, the border of a landscape, etc.)
Subjectivity	when geometry data are defined by adding an amount of interpretation rather than pure facts (e.g. characters passing through a city on a route not mentioned)
Vagueness	when poor definitions of a class of objects or individual objects do not allow a clear distinction of geometry data (e.g. a house downtown, as well as a house in a specific country)

Fictional places often tend to display all five of these criteria. On the one hand, scholars are faced with imprecise descriptions, so that they are forced to define one

possible geometry out of the imagined area. On the other hand, when it comes to the actual digitising, each person would obviously draw the geometry slightly differently. The results differ in scales of detail, more or less precise with a different idea of what could be the estimated extent of the described place of action. The results also depend on local knowledge and historical knowledge, as well as on personal preferences or skills.

In the research of literary geography, this is not at all a counteracting method. Quite the contrary: uncertain geometries matching the characterisation of the fictional space. As Piatti et al. (2009b) states: "This [formalising act of interpretation] is *not* what critics usually do. Instead they produce so-called secondary literature, interpretations and comments covering various aspects in works of fiction […] usually carefully balancing between several possible streams of interpretation. There, *ambiguity* is a sign of quality, here, in developing a database-supported literary cartography, it becomes a problem". Nevertheless, to have a comparable dataset concerning the size, the extent and the level of accuracy of a setting, there needs to be an *adequate data model,* that can deal with this heterogenic collection of data we are facing.

Having introduced the inherent property of uncertainty within fictional space, this paper will provide concrete solutions to enable spatial analysis and cartographic representation on the basis of a database-supported literary atlas information system. We start with some general declarations about the relation between textual space and geospace and the associated uncertainty. Subsequently, we continue with the actual data modelling. This is followed by a detailed characterisation of fictional spaces with the example of the categories "settings" and "routes" in order to introduce the necessary information to convert the described spaces into geodata. The paper will conclude with an outlook onto possible map- and analysis-outputs and plans for future work.

2 Terms and Definitions of Literary Theory

2.1 Geospace and Textual Space

In order to understand the requirements of the data model for an atlas of literature, an examination of the usage of space in both disciplines, the geosciences and the literary sciences, is necessary. Hereafter, we differentiate between **geospace** and **textual space**. Geospace is the real, physical space, which can be described geographically and represented cartographically, with measurable distances between any desired points. This space is called the *"first space"* (compare Soja 1996). In the contrast, textual space is the virtual space, generated and communicated by means of linguistic expressions. It may contain imaginative representations of the first space and functions as the "*second space*".

It is obvious that the specification of textual spaces, even if it refers to the same geographical extent, differs in every fiction. If one analyses the relation between

Table 1 Scale of reference classification (Extract of content-related attribute table, see Reuschel and Hurni 2011)

Relation Elations Between Textual Space and Geospace		
Terms	Explanations	Examples
Imported	The setting corresponds to geographical reality by realistically portraying the surrounding.	Libuše Moníková: *Verklärte Nacht* (1996): Moníková's characters stroll through Prague, each specific place is linked to a specific event or history. The reader is very well orientated within the town.
Transformed	The setting is transformed; this includes techniques of remodelling, renaming, relocation or a synthesis of several places.	K. von der Eider: *Antje Möller. Eine Eiderstedter Binnerdeern* (1911): Secondary literature identifies a lot of renaming and remodelling of settings within this story.
Invented	An invented setting within the known geographical reality.	Wilhelm Lobsien: *Der Halligpastor* (1914): One of the central settings of the story, a Hallig called "Westeroog" neighbouring to "Hoogeroog", does not exist in reality.
Imagined	There is no hint at all about the position of the setting, it is located "somewhere", with no real-world counterpart, only existing within the narration. Consequently, it is not possible to assign coordinates.	L. Frank Baum: *The Wonderful wizard of Oz* (1900): The "Land of Oz" is an autonomous geography surrounded by a desert, without any link to the real world.

textual space and geospace, it can be observed that literary plots range from the realistically rendered through various remodellings to the completely imaginary. In order to describe and formalise these relations, a four-part **scale of reference** is defined and used for places of actions (settings) and places of longing or remembering (projected places, see definitions at Sect. 3.1). Realistically portrayed places (*imported)* are arranged at the lower end of the scale, close to the geospace. All places that are somehow modified or alienated are assigned to the category *transformed*. Individual places that are purely fictional and integrated into the existing reality are covered by *invented*. In contrast, entirely *imaginary* spaces without any connection to our well-known reality are classified at the upper end of the scale, far from the geospace. Definitions and a selection of specific examples can be found in Table 1.

2.2 Scale of Reference and Uncertainty

At first sight, a relation between an increasing uncertainty and the scale of reference (Table 1) seems to be obvious: An imported setting that exists in reality should more easily locatable than transformed or even invented settings that are at a much higher level on the scale. Closer inspection shows, however, that a

correspondence between the scale of reference and the spatial uncertainty will be a fallacy. In fact, an author may include an invented place so cleverly and precisely into the existing environment that only a person with local knowledge would recognise it. Conversely, it may be that real, imported settings are only described superficially, or the author just mentions an approximate region. In those cases, the delineation possibilities of the corresponding places are rather uncertain. However, if the place of action is completely detached from reality, we reach the end of both scales: imaginary places are neither locatable nor representable on the basis of a topographic map.

Despite this categorisation, one has to be aware, that a certain geospace, used as textual space within different fictions, is never the same: Prague through the eyes of Paul Leppin (German Prague writer) is another Prague than through the eyes of Alois Jirásek (Czech Prague writer), even though both were living at the same time and using imported places by the majority. Their descriptions of Prague differ to such an extent that one might think they were writing about different towns. Prague is described as bizarre town, facing the past, in the works of German writers, whereas the Czech Prague is understood as a lively, booming city. On the one hand, the authors choose different literary topics (love story, crime fiction, historical fiction, utopic-fantastic fiction, etc.) or ambiance (for instant mythic, interchangeable scenery or symbolically-allegorically). On the other hand, time plays an important role. Both the time in which the story is set and the time in which the author is living influence the textual space. To put it differently: the authors' temporal and spatial perspective, their imagination and the changes of the geospace influence textual space. Many textual places are renamed or even do not exist anymore in the meantime. For these reasons, it is sometimes difficult to reconstruct the reality and to clearly classify every textual space into the scale of reference. The complexity of textual space is obvious: to characterise and visualise literary space scientifically, one needs more than just a geometry and a spatial category. This is also emphasized by Ungern-Sternberg (2009) in his work about mapping literature and narrations. According to him, "it is less important that a text names New Orleans than what it makes out of it". Joliveau (2009), who researches the relation of real and imaginary places in general, gains similar insights: "The spatial dimension of a document cannot be reduced to the juxta-position of place names. As demonstrated by Moretti [(1998, 2005)], it must also integrate elements of polarity, distances, itineraries and points of view".

3 Data Modelling for the Literary Atlas of Europe

3.1 Organisation of Fictional Space

When it comes to the actual analysis of textual space, we call it **fictional space**. This narratological generic term combines functions of all spatial objects.

A distinction is made between five spatial objects: *settings, projected places*, *markers*, and *routes* and *zones of action* (either of settings or projected places).

Setting	where the action takes place, characters are actually present
Projected Place	characters are not present, but they dream of, remember, or long for this place
Zones of Action	several settings or projected places combined to form a zone
Marker	a place that is just mentioned, but not part of the categories above; markers indicate the geographical range and horizon of a fictional space
Route	along which characters move through the fictional space; connections between waypoints with projected characteristics or setting's characteristics

Figure 1 illustrates schematically, how literary scholars differentiate between objects in fictional space. The spatial description and the existence of a (main) character and actual action are related to each other: if there is an action but the protagonist is dreaming or remembering it, it happens most probably on a projected place (middle column in Fig. 1). In this case, they have to decide to define an individual projected place or a zone with combined projected places or a projected route along the character is moving within its minds. The presence of action and the presence of characters are mostly directly linked to each other. Settings, projected places, routes and zones of action are inseparably connected with the story line, whereas markers merely serve as geographical indications.

However, it is not always possible to unambiguously determine a spatial description by one of the spatial categories. What are we going to do, for example, with jumps on the time scale: for instance, a fiction starts with a short frame story wherein the narrator remembers a past situation whereupon the actual story starts set in the past, again with the acting, present protagonist. Does the narrator's memory have the character of projected places or are those spatial descriptions of settings? Function changes (a place first marked as a topographical marker may become a setting later in the plot) or gradual changes of the described space itself (the changing of geospace may become the subject of the story) are, in fact, a different challenge. Therefore, in order to enable various interesting spatial analyses about literary space and to offer sophisticated visualisations of various properties, a number of attributes are added to each category.

3.2 Data Model

A data model was developed to realise and organise a coherent data basis for the 'Literary Atlas of Europe'. The data basis can be structured into four parts: general text information, including bibliography and assigned model region, data about the author, the temporal structure of the story line, and last but not least the spatial objects. All parts were taken into account within the data scheme shown in

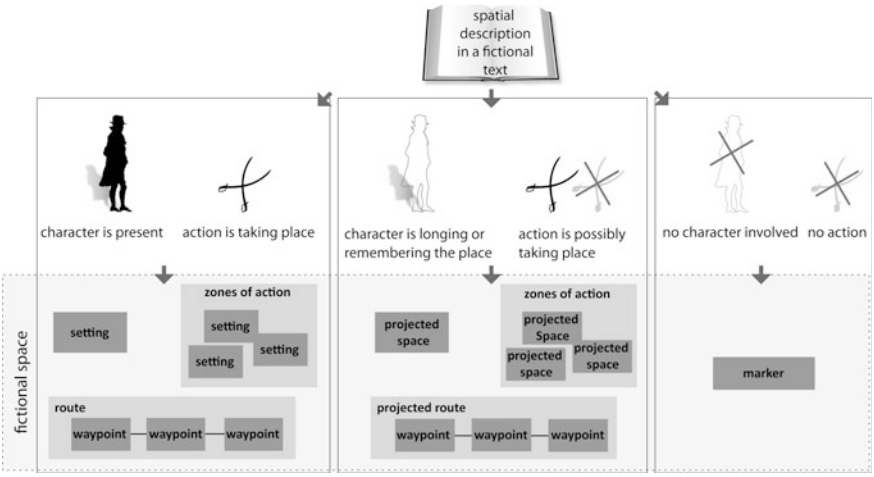

Fig. 1 Schematical differentiation of the fictional space

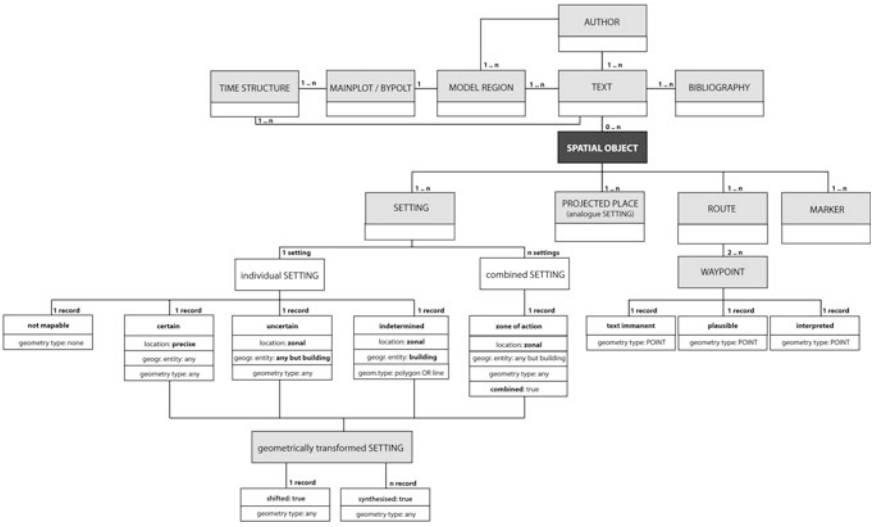

Fig. 2 Data model breaks down the space of a literary fiction into individual spatial objects

an overview in Fig. 2. Particular emphasis was put on modelling of the spatial structure of the text—the modelling of spatial objects. You will find the above-defined spatial entities that compose the geography of fiction and its complex structure. The geographical information is based on the Simple Features Specification (Open Geospatial Consortium 2011). Settings and projected places with uncertain, transformed or hardly locatable spatial objects necessitate

additional attributes and composed geometries respectively in order to reflect the fictional world. This leads to complex attributed spatial data. In Sect. 4 we will go more into detail and follow up with the data model of the categories "setting" and "route". The data model is implemented within a relational database management system and connected to an online submission form used to receive scholars' analyses of the data. Moreover, it serves as basis to process the automatically rendered maps for the 'Literary Atlas of Europe'.

3.3 Data Acquisition

To allow a comparative spatial analysis of literary space, its complex properties need to be broken down into geometries and attributes. This data preparation process involves careful interpretation, semantical and geometrical generalisation. We determined and predefined a number of attributes, which are assigned to individual spatial entities. Despite extensive data acquisition, it is impossible to reach complete representation through the transformation of spatial descriptions to geodata. We are aware of this, as Piatti et al. (2009a) advise: "be prepared to lose something—the complexity of a fictional text you encounter while performing a close reading (we lose style, vocabulary, psychology to name just a few elements)—but you also gain new insights." Furthermore, our chosen properties are not the only way to describe settings nor do we raise a claim for completeness. Quite the contrary: depending on the topic of the fictions, the framing research questions or the chosen geospace, attributes need to be supplemented or replaced.

As it is the main objective of the project to develop visualisations for literary spaces and to enable detailed spatial analysis, the question of the geometry acquisition arises. What possibilities exist to achieve geometries for literary spaces? An automatic text analysis of toponyms with a subsequent connection to an existing gazetteer database was ab initio not worth considering. Those techniques would not provide the desired complex and detailed information that is needed, such as coordinates from approximated zones or historical regions and would not find hidden, camouflaged places, not to mention additional thematic information. To capture the geographical locations of literary space and place, literary scholars are able to generate these themselves with the help of geocoding tools built upon Openstreetmap or Google Earth maps and satellite images, which are connected to the entry form (Fig. 3). In that way we have the possibility of covering worldwide geospace information. But the freely selectable scale of digitising leads to heterogenous data concerning scale, geometry type (point, line, polygon) and generalisation, respectively.

To enable extensive data acquisition, an online data submission form was developed, specially tailored to the needs of literary scholars. This submission form is connected directly to the database and controls the data reading and recording process of the scholars. It is designed to allow intermediate storage, data modification and manages redundant entries. For instance, a new text from an

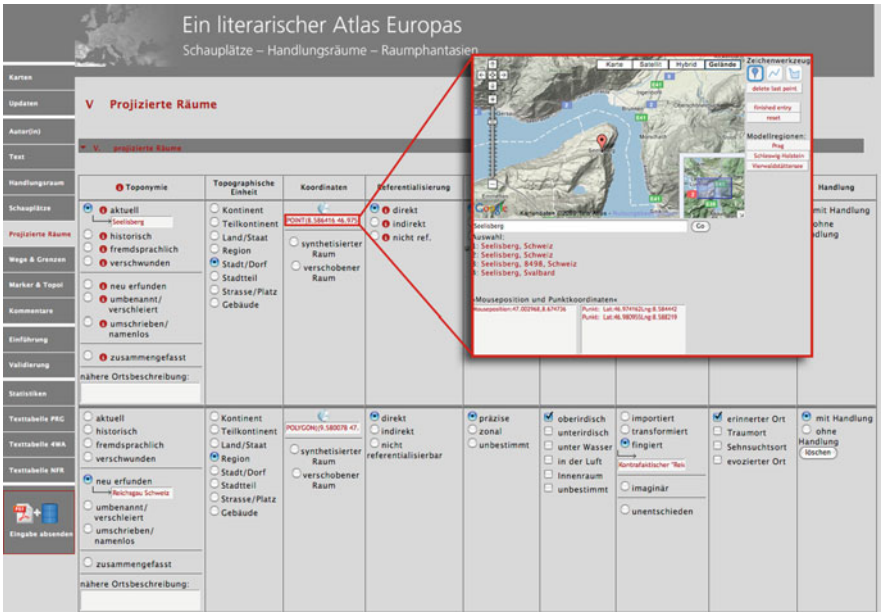

Fig. 3 Online submission form with digitising tool

already entered author is added without re-entering those bibliographic references, by assigning the fiction to the existing author record. Furthermore, the submission form provides scholars with an overview through "action-reaction" buttons: particular entry fields only appear when a specific button is pressed. To ensure an error-free entry, fields expecting numbers are marked and tested before sending it to the database. Similar precautions have been taken regarding coordinate fields: the coordinates are directly constructed through the geocoding tool. All data entered into the form are automatically evaluated and transferred into the database.

4 Concept and Realisation of Spatial Objects

4.1 Settings

Settings are defined as places of action where the fictional characters are present and acting. The way settings are described within a novel are quite different from fictional account to fictional account and have their own, inherent rules. Reuschel and Hurni (2011) identify five aspects (point 1. to 5.) that make fictional space unique and consequently so difficult to capture and visualise. Those aspects were developed for fictional space as a whole, but can easily be transferred to the main

category, the settings. Some further characteristics, which are specific for settings are additionally added (point 6. to 8.). The following list summarises them:

1. Settings are fragmentary, fictional space is completed and developed through the imagination of the reader.
2. Settings have uncertain, vague boundaries, neither physical nor natural, nor administrative, man-made boundaries.
3. Settings are sometimes difficult to localise and can result in an indeterminate location.
4. The "real-world counterpart" of a setting can lie in any time epoch; the intervening years may have altered the topography.
5. Settings can be transformed or remodelled by the author from the beginning or in the course of the story.
6. The extent of a setting can vary from a room inside a building to an international or worldwide level (theoretically, even galactic expansions such as in most science fiction).
7. The density of settings per text can vary between occasional and accumulated places.
8. Settings are used to create an ambiance or are used as a metaphor; they also have aesthetic functions.

Spatial analysis and visualisation of settings rely on the presence of coordinates. But using point coordinates only to spatially describe the location of a setting is not appropriate for a detailed spatial analysis, because the expansion of a place of action can indeed be determined in more detail. Depending on the map scale used for the acquisition and the extent of the setting, digitising with simple geocoding tools is—from the cartographic perspective—unattractive and leads to heterogenous data concerning scale, geometry type (point, line, polygon) and generalisation. A test analysis with students demonstrates how the same settings may be captured geometrically in different ways.

About ten groups of students were asked to analyse Schiller's *Wilhelm Tell* (1804) with settings in central Switzerland. An extract of the resulting geometries can be seen in Fig. 4.

To evaluate the results we compared them with each other. Three neighbouring villages are places of action. Not surprisingly, the students used either points or polygons with varying degrees of detail. This is also influenced by the base map information. With one single exception, all of the depicted geometries do overlap, even though they are depicted quite differently. Spatial statistical analysis is therefore possible. The visualisation of individual settings may show unpleasant rectangular shaped polygons, which should be avoided. We became aware of that no polygon definitions for villages really match their idea as fictional places. For instance, it would not help to use official data, such as those from swisstopo, the Swiss Federal Office of Topography. There, you would find built areas combining several villages. The official commune boundaries are defined politically and do not give the right impression of a spatial fictional description either. What would

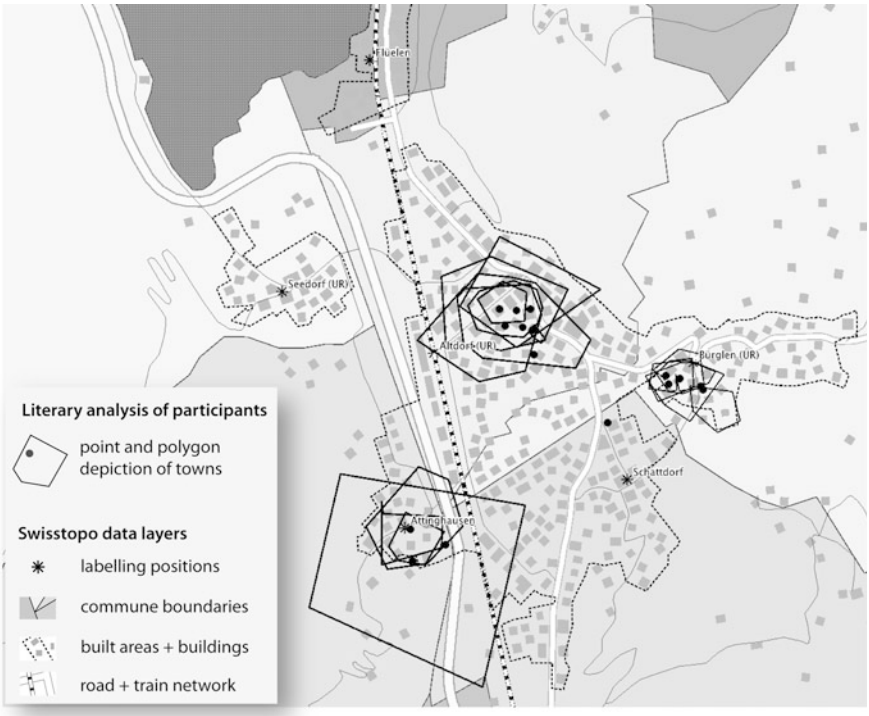

Fig. 4 Evaluation of town depictions of Friedrich Schiller: Wilhelm Tell. Literary analysis carried out by students

fit best at a medium scale within the model region of Lake Lucerne is labelling the position of villages.

To simplify the capture of geometry, we standardised our approach to digitising settings. Those within one of our model regions should be digitised with as much detail as possible. Since we work with three different model regions that vary in scale, the smallest possible geographical entity to be captured is determined as a point relative to scale: villages within the model region "Lake Lucerne" and "Northern Frisia", and buildings for the urban model region "Prague". This means that if, within Lucerne, there are a couple of buildings found to be settings, they are grouped within a zone of action and captured by a point (compare Fig. 5 left). Within Prague, however, each of the buildings is given its own geometry (compare Fig. 5 right). Settings that are located outside the model region are captured as points. In order to distinguish coordinate points that refer to a country, a village or just a building, we additionally ask for an indication of the approximated geographical entity. This information is used to present the entities on the map with different symbols and depending on the scale. Figure 5 schematically shows the usage of geometry types and the content per map scale respectively.

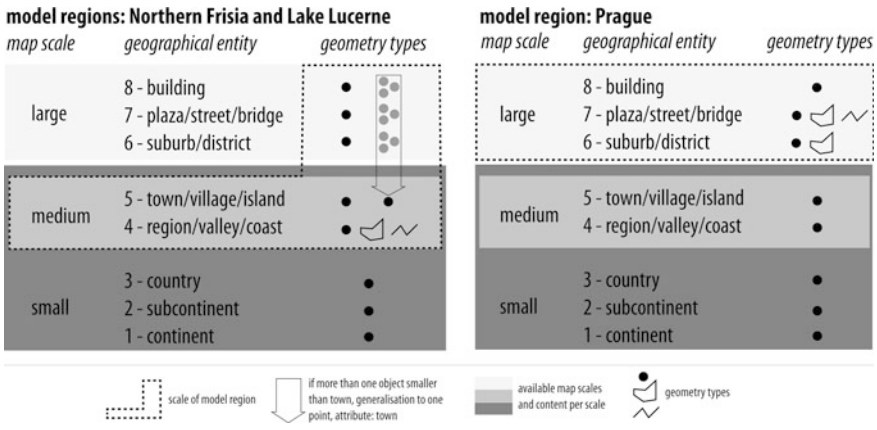

Fig. 5 Determination of the usage of geometry types and map content within and outside the model region. Medium scale model regions: Northern Frisia and Lake Lucerne (*left side*). Large scale model region: Prague (*right side*)

Other attributes used to visualise literary settings are a distinction between certain (precise) and uncertain (zonal) location and a specification of reference to actual geospace: imported, transformed, invented or imagined (Sect. 2.1). In addition, scholars can always add a comment or a quote to emphasise an interesting spatial feature within the text. With transformed places, there are two exceptional cases which we termed *shifted* and *synthesised settings*. In both cases, additional coordinates are necessary: a shifted setting needs one more geometry, the position of the original location (secondary geometry), whereas synthesised settings need at least one more geometry to break down the synthesis of the newly created place. In both cases, the primary geometry specifies the setting's position used within the fiction.

Another composed attribute is the toponym, or name of a setting. If the toponym used in the fiction differs from the current name or is not mentioned directly, we distinguish between a numbers of revealing properties, such as: *historical, foreign language, disappeared, neologism, renamed, veiled, paraphrased* or remaining *nameless.* If possible, the current name to which the toponym refers is also added. Consequently, three pieces of information are needed to specify the setting's designation: two different names and the property of the name used within the fiction. Visualisations of literary toponyms are discussed in Reuschel et al. (2009).

Further attributes of settings are used to describe the function and meaning of literary space. These can be summarised as *thematic attributes*. In contrast to *visualisation attributes*, thematic attributes do not describe the "where" and "what" of a setting; instead they provide interesting metaphors or are used for a particular purpose. Thematic attributes are only added if explicit indications are found in the text; they do not record interpretation or information from secondary literature. Table 2 lists and defines attributes that are used to thematically

Table 2 Setting attributes describing function and meaning of literary space

Function and meaning of literary space	
Simple scenery	The scenery is theoretically replaceable, it only serves as background, and it doesn't play an important role for the story.
Thematic scenery	The scenery is essential for the plot and is described plastically (portrayal of landscapes, architecture and atmosphere). Thus, the scenery is not replaceable.
Protagonistical physical	The setting takes physical action, usually in terms of natural phenomena or natural hazard events like avalanches, landslides or floods.
Protagonistical psychological	The setting influences the mind of the characters, for example by causing a feeling of being trapped or threatened; also positive feelings like freedom are possible.
Mythical connotation	The setting has obvious mythical ambiance or meaning.
Allegorical, symbolical connotation	The setting has an obvious metaphoric or symbolic connotation.

characterise settings. There are also some thematical attributes related to each model region, but due to particular local or historical knowledge we have not listed those attributes too.

Let us take a closer look at the implementation of all properties into the settings data model at Fig. 6. We distinguish between combined settings and individual settings. The latter can further categorised as certain, uncertain and undetermined objects. Except in those cases where a setting remains unmappable, all of the objects contain one simple feature geometry and describing attributes. Unmappable settings only contain attributes but cannot be used for a spatial analysis because of the missing geometry. If a geometric transformation has been determined, a secondary geometry can be added as separate object. In contrast to usual geodata models, where attribute values are tied to each individual entity, literary attributes of interest are not necessarily connected to the setting object, but rather to the text. This happens if a property—for instance, a mythical connotation (compare Table 2)—cannot be attached to a particular spatial object, but after reading the whole text, there are nevertheless clear signs of a mythical connotation within the text. In this case, any spatial object inherits the property from the text object.

As mentioned, necessary attributes that describe a setting can be divided into visualisation attributes and thematical attributes. In the following, three selected settings, one from each model region, are intended to show the diversity of the category setting and the way of obtaining and handling information for both visualiation and thematic attributes. The first example, Fig. 7, is taken from a novel called *Quatemberkinder* written by the Swiss author Tim Krohn. In his novel, the mythical world of legendary creatures is merged with the human world. Thus it is possible to directly link mythical connotations to the novel's settings. In this case, an additional comment provides information about what is happening. Krohn describes the settings of his homeland very precisely and in depth. In an

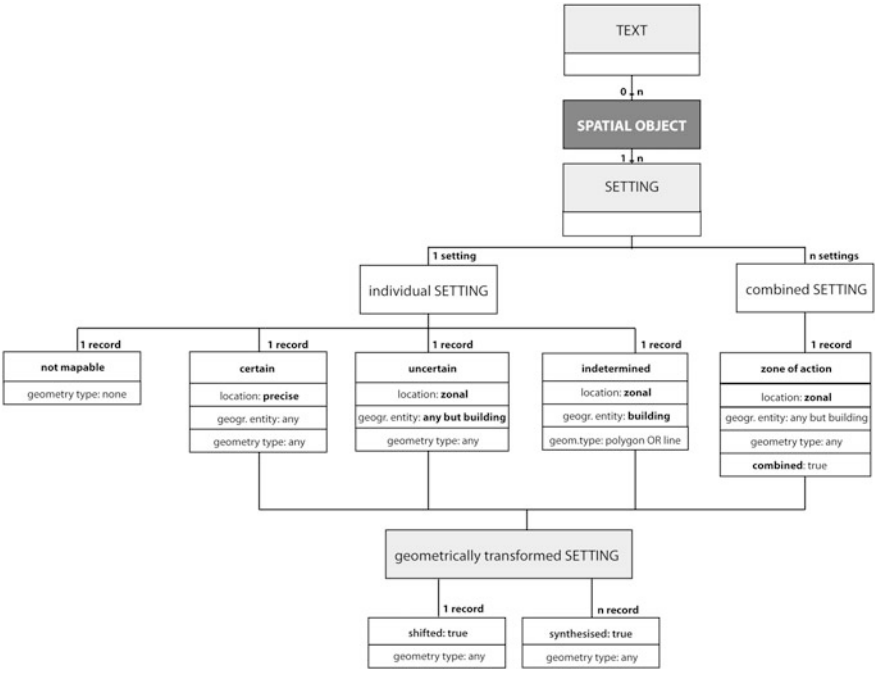

Fig. 6 *Data modelling of the literary spatial object: Setting*

interview he stated: "Writing the novel, I orientated myself with maps and a guide book. I waived doing on-site research, because it would have limited my opportunities" (Geisel 2007). Appropriate attributes for the setting were entered as in Fig. 7. Lake Zug is roughly digitised by a polygon and visualised as a red area with a blurred border, indicating a realistic imported place.

The second example is from the analysis of Paul Leppins ghost novel *Severins Gang in die Finsternis* (English translation: Severin's Journey into the Dark) of 1914. The scenery of Prague is inseparably linked to his novel. Like almost all German-speaking writers from Prague, he describes a bizarre town that comes to a standstill, facing the past. Within their texts, they revive legends and use locations such as jumble markets, cemeteries or morbid houses. The digitised polygon in Fig. 8 is the zonal position of a building, Dr. Konrad's atelier. The resulting representation shows a labelled symbol in the centre of the polygon, indicated by thin radial lines starting from the centre to the polygon's border.

The third and last example (Fig. 9) is taken from a Frisian novel, *Rüm haart—klar kimming* (Frisian slogan translated with "open heart—clear horizon"), written by Thusnelda Kühl. In terms of literary geography, this fiction from 1903 has a number of interesting settings. The one chosen for this example describes a ferry crossing that lands directly next to a restaurant. This river crossing is used as a metaphor for the transition between village idyll and the outside world. However,

Setting:

Attribute	Value
Toponym used in fiction	Zugersee (Lake Zug)
Location	precise
Location details	above-ground
Reference to reality	imported
Direct reference	true
Geographical entity	region
Mythical connotation	true
▼ Comment to mythical connotation	On this place a romance between the protagonist and a mermaid takes place. Mythical creatures exist naturally beside humans. Two different worlds do exist and interact with each other.

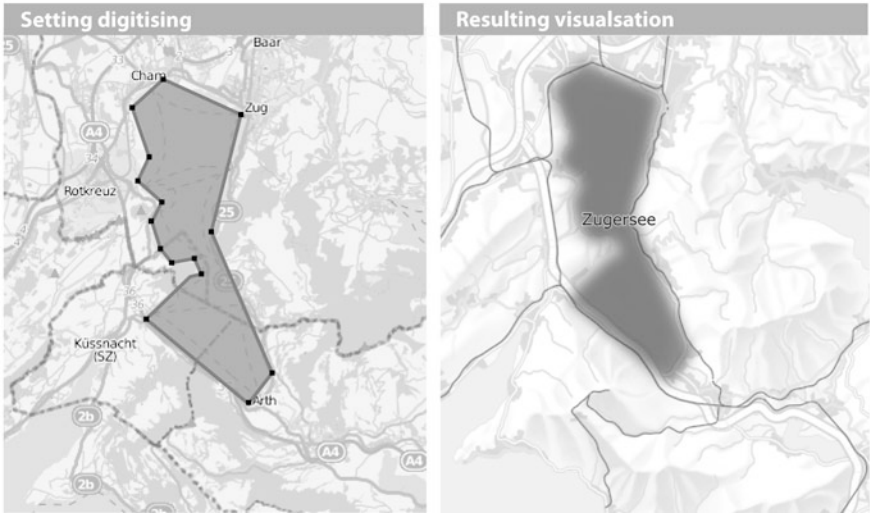

Fig. 7 Complete depiction of a setting, model region of Lake Lucerne, taken from Tim Krohn: Quatemberkinder. (Literary text analysis carried out by B. Piatti)

inquiries revealed that there is no ferry at this point. Kühl was apparently relocating an existing ferry service from down the river for her purposes. The database needs two separate geometries to record this shift: the primary position from the author's description and a secondary, the real position of the ferry service. In this way, the direction can be represented with a dashed Bezier curve. Thereby only the primary geometry is colour-coded as a setting, whereas the secondary geometry remains grey, like the base map, the "real world model".

Setting:

Attribute	Value
Toponym used in fiction	Dr. Konrads Atelier
Location	zonal
▼ Comment to location	This setting is located in one of the new buildings of the redevelopped area. Consequently it is located somewhere in Josefov, the former Jewish quarter.
Location details	indoor
Reference to reality	imported
Direct reference	true
Geographical entity	building
Thematic scenery	true
▼ Comment to thematic scenery	The scenery of the redevelopped area is essential for the plot and not replaceable.

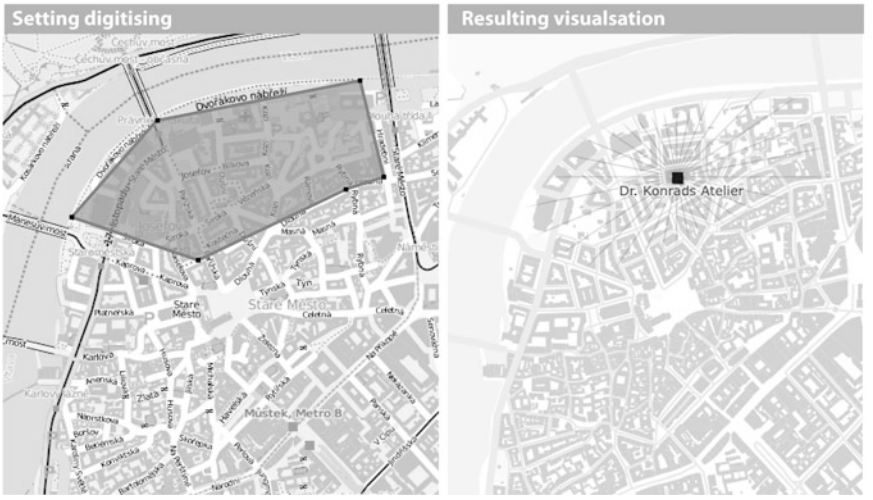

Fig. 8 Complete depiction of a setting, model region of Prague, taken from Paul Leppin: Severin's Journey into the Dark. (Literary text analysis carried out by E. Markvartová)

4.2 Routes

One of the most challenging categories within literary geography are routes, the dynamic element that links settings or projected places to each other. Here, vague, uncertain locations meet vague or even unknown connections. We distinguish between implicit and explicit routes. Implicit routes do not thematise the course of a journey; they are comparable with a movie cut: characters leave place A and the

Setting:

Attribute	Value
Toponym used in fiction	Fähranleger (ferry pier)
Toponym description	nameless, described
Location	precise
Location details	above-ground
Reference to reality	transformed
Translation	true
▼ Comment to translation	translated: the author placed the pier directly next to the restaurant called 'Fährkrug'.
indirect reference	true: secondary literature
▼ Comment to indirect reference	there is no ferry at 'Rothenspieker habour', it had to be the ferry service between 'Wollersum habour' and 'Oldenswater Vorland' down the river
Geographical entity	plaza / street / bridge
Thematic scenery	true
▼ Comment to thematic scenery	The river crossing functions as transition between village idyll and the 'outside world'.

Fig. 9 Complete depiction of a setting, model region of Northern Frisia, taken from Thusnelda Kühl: Rüm haart—klar kimming. (Literary text analysis carried out by K. Seifert and K. Winkler)

next moment arrive at place B. However, explicit routes describe the actual movement of the characters and the places they are passing. Within the scope of our project we are only interested in explicit routes. Nevertheless, there is variation in the completeness and comprehensiveness of explicit routes. In some cases, routes can be used as detailed travel guides, while often the route network remains patchy and imprecise. Similar to settings, we defined specific characteristics that are inherent to routes and need to be considered during the process of data modelling and visualisation. These are:

1. Start or end of a route is often untraceable: characters show up "out of nowhere", or their path dissipates "into nowhere".
2. Route linkages are often ambiguous; drawing connections requires interpretation.
3. The scale may change within a route description.
4. Routes are described by individual toponyms along the way.

To depict literary routes in as much detail as possible, an acquisition in the form of a line is insufficient. Rather it requires additional attributes to waypoints, which comprise the individual elements of routes. The start and end point of each route is differentiated between "known" and "unknown". In the case of "unknown", the waypoint is the first or last known location before the characters (dis)appear. To distinguish between waypoints that can be taken directly from the text and those that result from interpretation, we classify each waypoint into one of three categories: **text immanent point** (taken directly from the text), **plausible points** (anchor points) and **interpreted points** (the most probable way according to the scholar's analysis). Text immanent points are also provided with a toponym. Furthermore, the information of the moving character is attached to each complete route construct, whether or not the route is projected. This individual information enables different ways of handling and visualising the category of routes. If one is only interested in facts from the fiction, one chooses a **schematised route** that only consists of text immanent points. The alternative is an **interpreted route** including all kinds of waypoints. Note that a route is defined as a connection between waypoints with projected properties or setting's properties, but routes are of different degree of details than the spatial object setting or projected place. This means, if an intermediate place of a route is a setting or a projected place simultaneously, an additional more detailed entry of the setting or projected place is required.

Let us take a closer look at the routes data model in Fig. 10. The route object itself does not contain any geometries, only the general attributes mentioned above. The attributed point geometries are part of the waypoint object. At least two waypoint objects form a route object. Waypoint objects with different attributes enable alternative route creations.

A detailed example from a spatial analysis of Paul Verne's *De Rotterdam à Copenhague à bord du yacht à vapeur Saint-Michel* (1881) will demonstrate the acquisition of data for routes. The text is about a common ship journey taken by the Verne brothers, Paul and Jules. The title of the text already hints at the

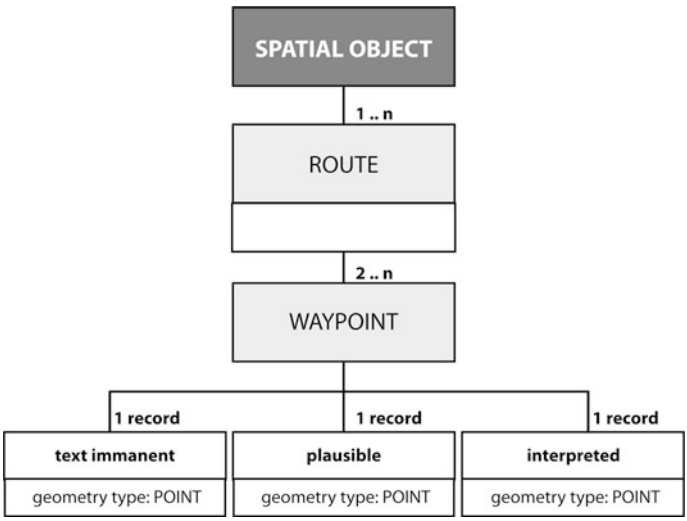

Fig. 10 Data modelling of the literary spatial object: route

departure harbour and the port of destination. Several intervening stations, where names are known and mentioned in the text, were added as text immanent points. All other points, not directly mentioned, could be added plausibly because the means of transportation is by water. Remarkable is a changing of the route's level of detail: Sections along the North Sea coast were digitised quite generously, from harbour to harbour, probably from a small-scale map (outside Fig. 11), whilst the route along the meandering Kiel Canal within the model region of Northern Frisia is determined in too much detail to be shown on the medium scale map. With the help of the point attributes, two types of routes are derived as shown in the resulting visualisation. The schematised route (Fig. 11, *middle*) connects all settings along the cruise, whereas the interpreted route (Fig. 11, *bottom*) gives a quite detailed impression of the travel.

5 Summary and Outlook

The data model we have developed for literary spatiality provides a basis for the 'Literary Atlas of Europe'. Our approach includes the usage of attributed geometries to deal with inherent uncertainty properties, as demonstrated at the outset. The atlas will allow scholars to see patterns and information that are literally invisible. It will work as a generator for new ideas and may provide foundations to answer questions such as: *When does the fictionalisation process of a certain region start and are there still people writing about this region today? Are there particular regions that are often mythical connoted?* You can even think of more complex questions like: *Under which (political-historical and other) conditions*

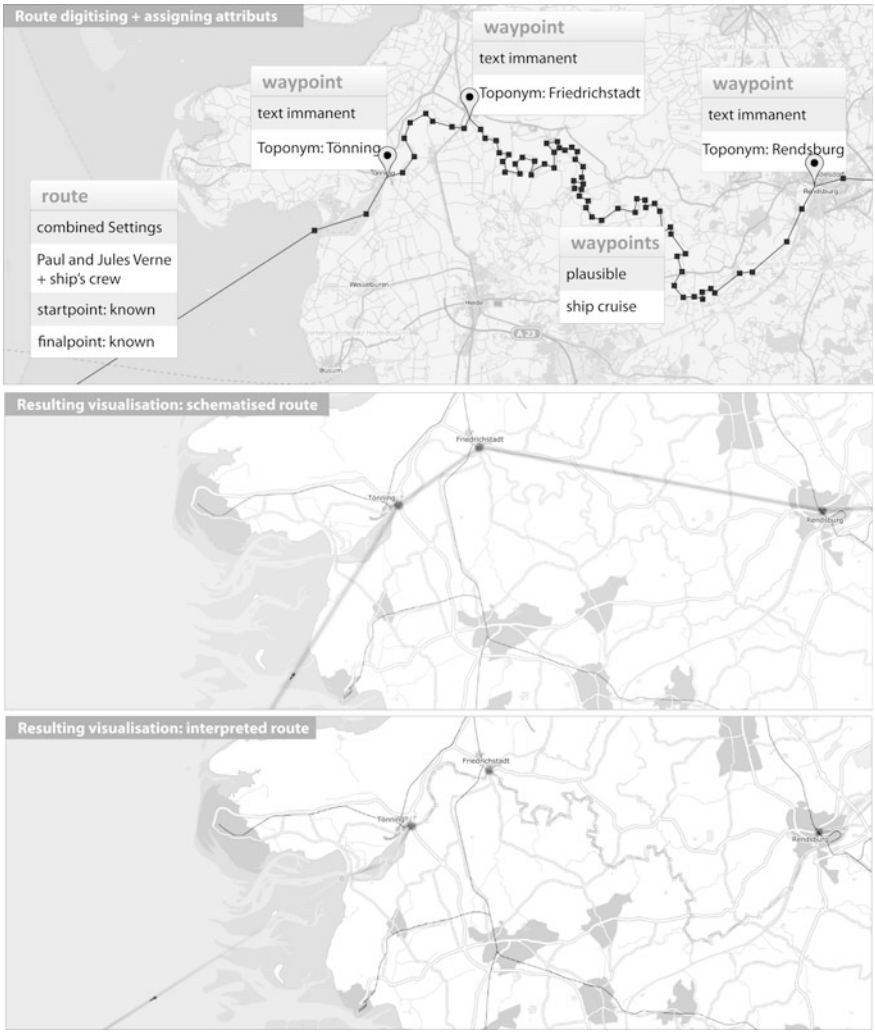

Fig. 11 Complete depiction of a route, model region of Northern Frisia, taken from Paul Verne: De Rotterdam à Copenhague à bord du yacht à vapeur Saint-Michel. (Literary text analysis carried out by K. Seifert and K. Winkler)

does the (imagination-) space of literature contract, and under which does it expand? How international is the space, or do almost exclusively native authors describe it? The generated maps are thereby the visual basis, interactive tool and output of the spatial analysis of literature. The subsequent task for literary scholars is to determine the reasons and connections for the resulting patterns. The database-oriented approach enables spatial and thematic analysis. Various perspectives on fictional spaces can be achieved through almost unlimited possibilities of

spatial data combinations controlled by their attributes. The data model allows us to show how literature creates and organises space, for entirely different fictional worlds, with different elements and density.

In order to visually analyse the results on a map, two different map types are currently under development. On the one side, we visualise individual spatial objects with a scientific uncertainty geovisualisation, for instance, for a single text. On the other side, we calculate statistical surfaces of a statistically significant number of literary places, resulting from queries to the whole data corpus. First results can already be found in Reuschel and Hurni (2011), Bär and Hurni (2011). Both automated visualisations are working with the data model presented here.

Further efforts are necessary to enhance the examination possibilities of a fictional space. Is it, for example, possible to show a detailed fictional space not only in a very enlarged map section, but also in an overall view? How do we handle places that change their functions within the text? We also investigate the integration of further properties of spatial objects that could be of interest for future literary work. For example, how could we add and visualise a chronological order of the important settings? Could we analyse from which place a projected place has been triggered? How can we model moods of different characters associated with a place of action into the existing model? The close cooperation between cartographers and literary scholars remains an indispensable prerequisite to achieve its ambition. So we agree and conclude with the insights of Joliveau (2009): "There is a lot of room for innovations in this domain, but we need to focus on improving the connection between researches in scientific geovisualization and artistic experimentations".

Acknowledgments This paper is part of the project 'Literary Atlas of Europe' and supported by the Swiss National Foundation, grant number CR21I2_130160.

References

Ahlqvist. O. (2009). Visualization of vague category counts—introducing the fuzzy dot density map. *Proceedings of the 24th International Cartographic Conference, Santiago de Chile, Chile*.

Bär, H. R., & Hurni, L. (2011). Improved density estimation for the visualisation of literary spaces. *The Cartographic Journal, 48*(4), 309–316.

Bennett, B. (2001). What is a forest? On the vagueness of certain geographic concepts. *Topoi, 20*(2), 189–201.

Bittner, T., & Stell, J. G. (2002). Vagueness and rough location. *Geoinformatica, 6*(2), 99–121.

Cohen, P. (2011). *Digital maps are giving scholars the historical lay of the land*. The New York Times, July 26, 2011. http://www.nytimes.com/2011/07/27/arts/geographic-information-systems-help-scholars-see-history.html?_r=1. Accessed 13 April 2012.

Drecki, I. (2007). Geographical information uncertainty: The concept and representational challenges. *Proceedings of the 23rd International Cartographic Conference, Moscow, Russia*.

Egenhofer, M., & Mark, D. (1995). Naive geography. COSIT'95, Semmering, Austria, In A. Frank & W. Kuhn (Eds.), *Lecture notes in computer science*, Vol. 988, pp. 1–15 Berlin: Springer.

Erwig, M., & Schneider, M. (1997). Vague Regions. In M. Scholl & A. Voisard (Eds.), *On advances in spatial databases, lecture notes in computer science*, Vol. 1262, pp. 298–320. Berlin: Springer.

Fisher, P. F. (1999). Models of uncertainty in spatial data. In *Geographical information systems—principles and technical issues* (pp. 191–205) New York: Wiley.

Geisel, S. (2007). *Bei den Seelenen*. Neue Zürcher Zeitung AG, Nov 6, 2007. http://www.nzz.ch/nachrichten/kultur/aktuell/bei_den_seelenen_1.580094.html Accessed 13 April 2012.

Griethe, H., & Schumann, H. (2006). The visualization of uncertain data: Methods and problems. *Proceedings SimVis'06, Magdeburg, Germany*.

Joliveau, T. (2009). Connection real and imaginary places through geospatial technologies: Examples from set-jetting and art-oriented tourism. *The Cartographic Journal, 46*(1), 36–45.

MacEachren, A. M, Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M.,& Hetzler, E. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science, 32*(3), pp. 139–160.

Montello, D. R., Goodchild, M. F., Gottsegen, J., Fohl, P. (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. In *Spatial Cognitions and Computation*, 3(2&3), pp. 185–204.

Moretti, F. (1998). *Atlas of the European Novel 1800–1900*. London: Verso.

Moretti, F. (2005). *Graphs, maps, trees. Abstract models for a literary history*. London: Verso.

Open Geospatial Consortium. (2011). *OpenGIS Implementation Specification for Geographic information—Simple feature access—Part 1: Common architecture*. http://www.opengeospatial.org/standards/sfa. Accessed 4 Nov 2011.

Pang, A. T. (2001). Visualizating uncertainty in geo-spatial data. In *Proceedings of the Workshop on the Intersections Between Geospatial Information and Information Technology. National Academies Committee of the Computer Science and Telecommunications Board, Washington, D.C.*

Piatti, B., Reuschel, A.-K., Bär, H. R., Cartwright, W., & Hurni, L. (2009a). Mapping literature. Towards a geography of fiction. In w. Cartwright et al. (Ed.) *Cartography and Art* (pp. 177–192) Berlin: Springer.

Piatti, B, Reuschel, A.K., & Hurni, L. (2009b). Literary geography—or how Cartographers open up a new dimension for literary studies. In: *Proceedings of the 24th International Cartographic Conference, Santiago de Chile, Chile*.

Piatti, B. (2008). *Die Geographie der Literatur. Schauplätze, Handlungsräume, Raumphantasien*. Göttingen: Wallstein-Verlag (424 pages; 2nd edition 2009).

Reuschel, A. K., Piatti, B., Hurni, L. (2009). Mapping literature. The prototype of "A Literary Atlas of Europe". In: *Proceedings of the 24th International Cartographic Conference, Santiago de Chile, Chile*.

Reuschel, A. K., & Hurni, L. (2011). Mapping literature: Visualisation of spatial uncertainty in fiction. *The Cartographic Journal, 48*(4), 293–308.

Soja, E. (1996). *Thirdspace, journeys to Los Angeles and other real-and-imagined places*. Oxford: Basil Blackwell.

Thomson, J., Hetzler, B., MacEachren, A. M., Gahegan, M. N., & Pavel, M. (2005). Typology for visualizing uncertainty. In: *Visualization and data analysis* (5669, pp. 146–157).

Ungern-Sternberg, A. (2009). Dots, lines, areas and words: Mapping literature and narration (With some remarks on Kate Chopin's "The Awakening"). In W. Cartwright et al. (Ed.) *Cartography and Art* (pp. 229–252) Berlin: Springer.

Waters, T., & Evans, A. J. (2003). Tools for web-based GIS mapping of a "fuzzy" vernacular geography In: *Proceedings of the 7th International Conference on Geocomputation*, 8–10 September 2003