

Linear referencing to assist the classification of natural and regulated Danish river systems

Thomas Balstrøm, Sebastian H. Mernild & Bent Hasholt
Institute of Geography, University of Copenhagen, Denmark.

E-mail address for corresponding author: tb@geogr.ku.dk

Abstract

To implement a classification system for natural and regulated Danish rivers a linear referencing system has been established for the county of Sønderjylland in southern Jutland, Denmark. Routes created from a map of rivers were defined and segments along the rivers' reaches were assigned the classification values of various geographic map units extracted from polygon maps (topsoil textural composition, subsurface geology, geomorphology and potential ochre pollution). In this way the river segments were subdivided and ready for explorations of presence/absence or co-occurrences of, say, biological activities in different physical environments.

Introduction and aims

In December 2000 the European Commission published the directive 'Establishing a Framework for Community Action in the field of Water Policy (Water Frame Directive). The background is to prevent the increasing deterioration and to protect and enhance the status of aquatic ecosystems. This initiative is one of many that focus upon the local river systems and their environmental state and underlines the need for registrations of hydrologic, morphodynamic and chemical information (quantitative as well as qualitative) about their outer and interior conditions.

Related hereto Hasholt (1999) suggested a system for classification and regionalisation of Danish watercourses (natural and regulated) based on physical parameters. This work was followed up by Mernild & Hasholt (2003) who performed a test of the suggested system from field investigations of 42 natural watercourses in different types of landscapes.

This paper describes a digital implementation of the linear referencing system (LRS) included in ESRIs ArcGIS 9.0 for the rivers within the county of Sønderjylland located in the southern part of Jutland, Denmark. The LRS is adopted to build a geodatabase of river information in order to stratify and sort the huge amount of information obtained when assembling the physical parameters from the landscapes. After completion the system may be used to perform searches for and visualize co-occurrences of various physical hydromorphologic conditions and to compare those with registrations of biologic assessments of water quality.

Description of the digital setup

Digital storage of river information in geographical information systems has been ongoing for years but, typically, it has been represented by point, line and polygon features according to abstractions and generalizations of real world phenomena. Such traditional data storage has formed a poor basis

for queries concerning cross-feature occurrences of observations along rivers and elements outside their reaches. For example, it has been difficult to analyze occurrences of bottom sediment types of river segments stored as lines with measurements from monitoring stations stored as points and information about the river segments' watersheds, the soil types etc. along the reaches stored as polygons.

Meanwhile, a so called linear referencing system (LRS) includes a set of components better suited for storage of information related to rivers. The LRS was originally created as a tool to keep track of highway information by defining a dynamic segmentation system describing the local conditions of highways concerning pavement material, age, quality, ownership, width, maximum allowed speed etc. (Figure 1). In hydrology applications, linear referencing may be used to store information from field monitoring stations along the rivers' reaches. In such connections the term river addressing is used more commonly (ESRI, 2004).

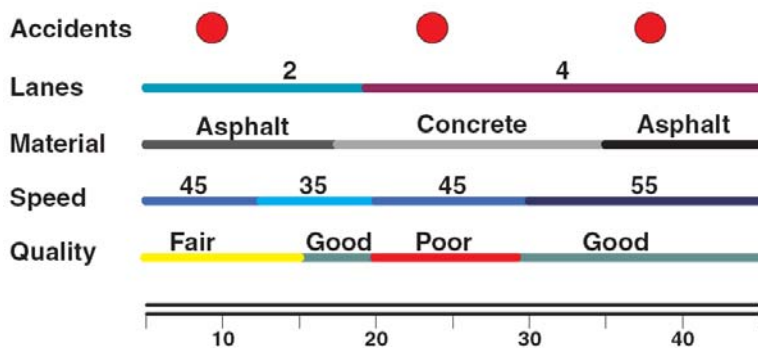


Figure 1 Storing multiple sets of attributes for one feature (in this case a highway) is possible using linear referencing.

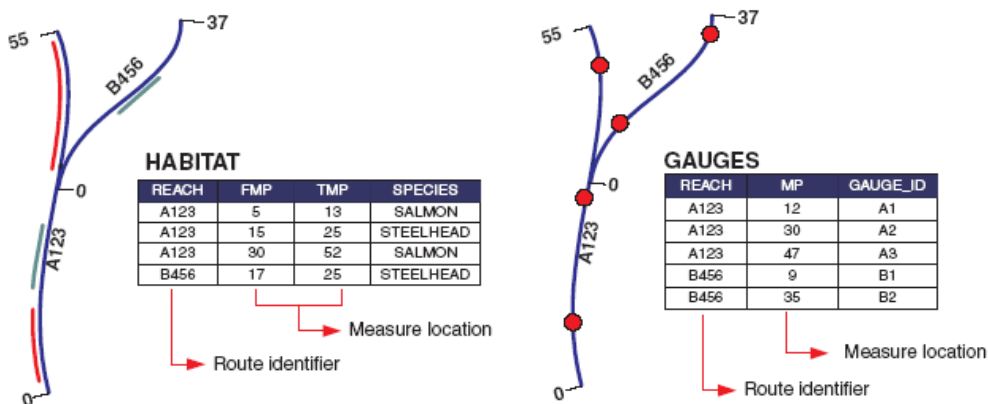


Figure 2 Example of line and point events located along a route. To the left is presented the location of fish habitats as linear events along river reaches. The FMP (from measure point) and TMP (to measure point) columns indicate at what distances relative to the origin of the reach where observations occur. At each junction the distance is (re)set to 0 for the next reach. To the right is presented the location of flow gauges as point events.

Establishment of the geodatabase

A personal geodatabase for the county of Soenderjylland about the rivers and the environmental features surrounding them has been established. The county of Soenderjylland was chosen because it represents a cross section of typical landforms in Jutland. The location of the study area is presented in Figure 3. As the base map for all streams and rivers the county has provided their line shapefile of the geometric river network and associated attributes which (among others) include a biologic assessment of each major river segment. Furthermore, it is an example of the actual information available in the administration of watercourses in Danish counties.

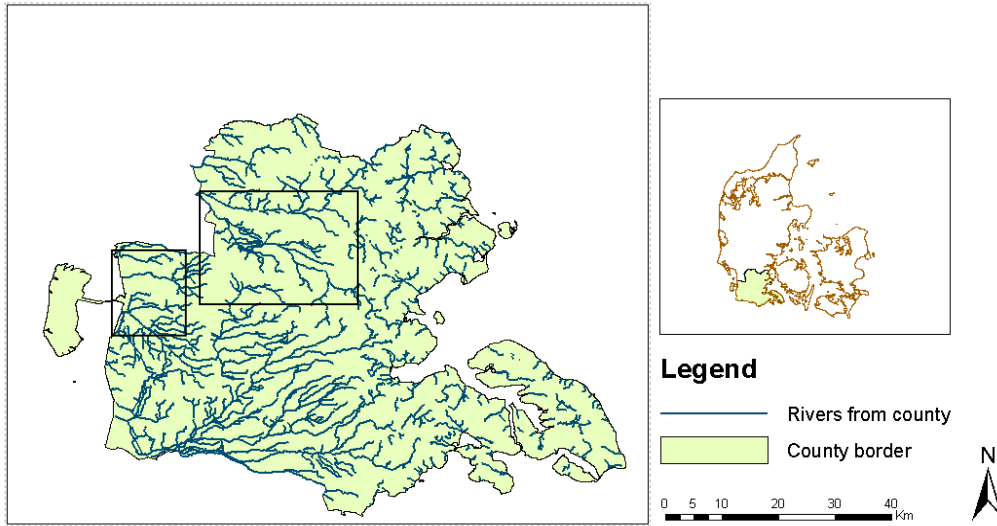


Figure 3 The location of rivers within Soenderjylland's County. Please note that rivers on the island of Roemoe in the westernmost part of the county are not included in the data set. The dimmed area in the westernmost part of the study area is presented in Figure 4. The dimmed area in the central part is presented in Figure 6.

This county's base map is, unfortunately, not compatible with the hydrologic theme available from the Danish National Survey and Cadastre at a scale of 1:10,000 but assembled from a mixture of various map sources collected 5-20 years ago. The sources are (mainly) Danish Institute of Agricultural Sciences (DIAS) and digitised analogue maps from Danish National Survey and Cadastre at a scale of 1:25,000.

A thorough investigation of the base map revealed the following inconsistencies:

1. it lacks the major river systems for the Island of Roemoe in the westernmost part of the county
2. it lacks all minor river systems
3. it includes several inaccuracies concerning the actual placement of the major river systems
4. it lacks connectivity of river segments in order to define the ordering of streams.

A visualization of the differences between the reference used and the newest river theme available from the Danish National Survey and Cadastre is presented in Figure 4. In spite of the problems mentioned above it was decided to use it as a reference because it includes the abovementioned biologic classification of the major rivers and expresses the actual state of the art regarding hydrologic data available for the county.

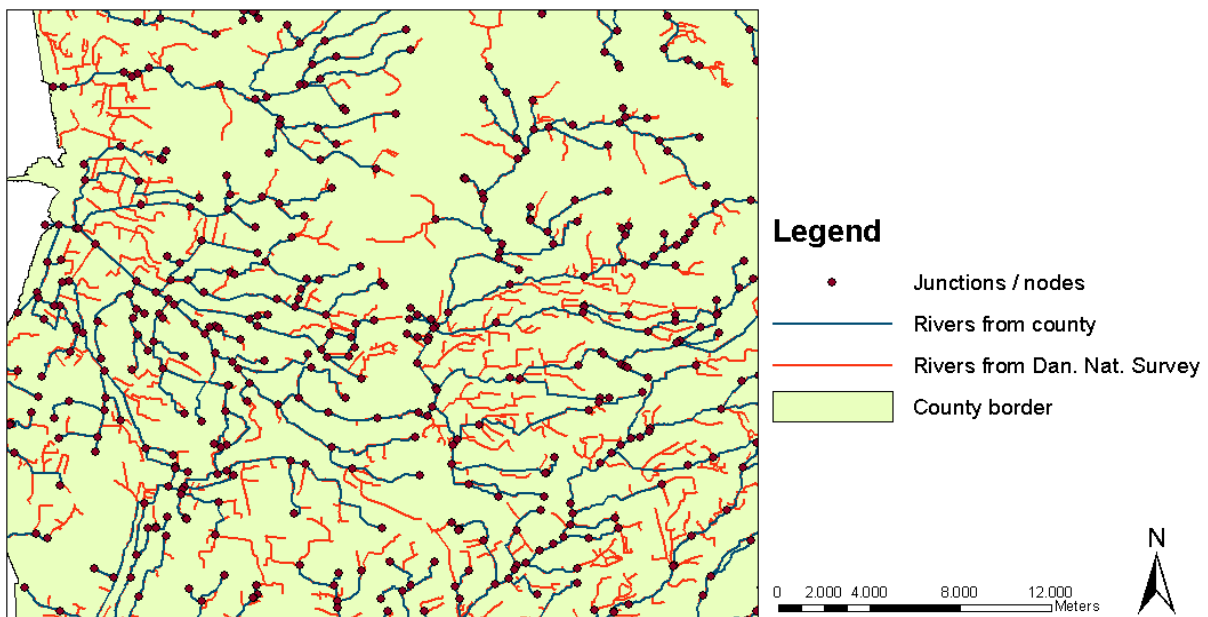


Figure 4 A zoom into the westernmost part of the county presented in Figure 1 and display of the junctions of rivers as compiled by the county and the rivers as registered by the Danish National Survey and Cadastre. As seen the data from the county includes the major river system, only, and certain discrepancies between the two data sets within the major river system occur as well.

To establish a linear referencing system the river polyline theme was used to define a system of *routes*. In this a unique identification of river segments located between junctions (nodes) is created and a measurement system stored for each of these segments is defined as well. When initialized the total number of river segments found is 1555. For each of them measures are build which are values defining distances along each river segment between junctions.

After this initialization the routes may be subdivided based on the location of polygons along the routes. Thus, a river's segments may be subdivided according to locations of, say, various geomorphologic features along its reach. This task is carried out by deriving a line event from each polygon found along the river's route reference.

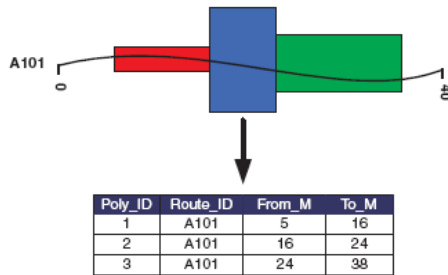


Figure 5 Polygon information is transferred to routes from computations of the geometric intersections of the route and the polygon data. The From_M and To_M indicates the from- and to-locations of the events relative to the starting point at 0.

Actually, this reminds of a classic overlay analysis where the locations of various polygons along a polyline are ‘stamped’ onto the polyline by physically splitting this into a number of new individual polylines where the polygons intersect the original polyline. However, using the LRS a physical split of the polylines does not occur. Instead, the intersections are stored in the measurement system as line events, only, that keeps track of the position of the ‘intersections’ or rather said where changes in a specific subtype occur along a route, see Figure 5. In this way it is possible to subdivide individual routes as required based on several individual polygon themes representing different geographic features.

Four polygon feature datasets were transferred to line events in the database:

1. Information about top soil texture divided into 8 categories found in the upper 0-30 cm of the soil on agricultural land. For non-agricultural areas the general land use was entered. Source: DIAS. Original scale: 1:50,000. 4692 events.
2. Information about the geological sediments found at a depth of 1m below the surface. Source: Danish Geological Survey. Original scale: 1:25,000. 7753 events.
3. Geomorphology. Source: DIAS, Danish Geological Survey. Original scale: 1:100,000. 2331 events.
4. Lowlands and potential ochre pollution. Source: DIAS. Original scale: 1:50,000. 5531 events.

When created the four line events were intersected one by one leading to one final event theme in which all information about top soil texture, geological subsurface type, geomorphology and occurrence of lowlands were combined. Thus, this event theme includes 15262 events. An example of a search in the database is presented in Figure 6 together with a map of the monitored environmental state of the watercourses.

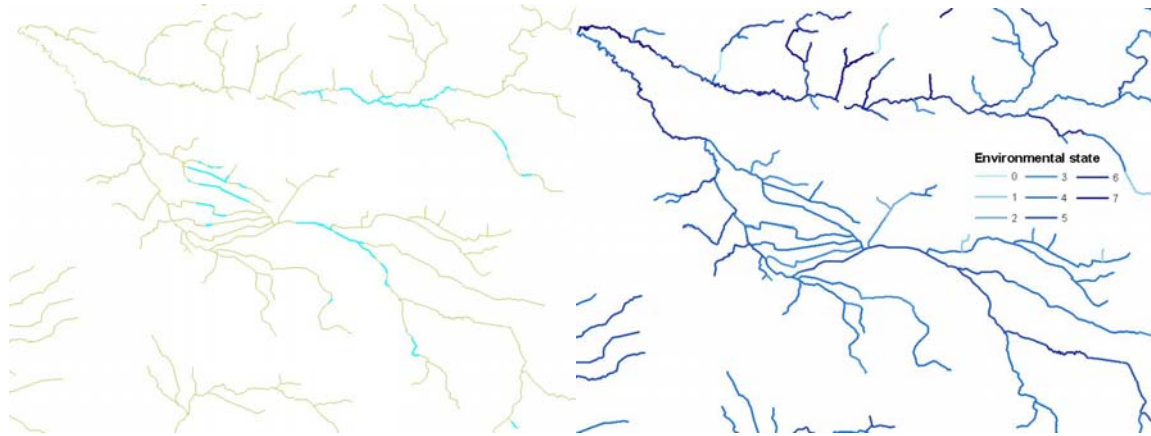


Figure 6 Left: The result for a search of line events dominated by a potential severe ochre pollution found on peaty deposits in top- and subsoils in the glacial landscape from the Weichsel era. The area showed has an extension of 30 x 25 km with its centre located at the village of Tiset (for the spatial location of the area please also refer to Figure 3). **Right:** The environmental state of the watercourses shown in the left figure. The reaches found have an environmental state in group 3-5 according to the County of Soenderjylland and guidelines from Miljø- og Energiministeriet (1998).

Although not yet implemented the final event theme could be joined with additional information about the watercourses terrestrial features such as the river segments' slope, sinuosity and biologic assessment. How these themes were constructed is described below.

The basin relief

The basin relief was derived from a digital elevation model (DEM) provided by the Danish National Survey and Cadastre with a cell resolution of 25 metres. Local basins were found using the fill, flowdirection and basin functions in ESRI's Spatial Analyst extension for ArcGIS and, subsequently, a zonalstatistics function found the range in elevations within each basin. Afterwards, a fifth line event table was calculated for the river polylines found within each basin from an assignment of the elevation range using the earlier mentioned LRS function 'locate features within polygon' tool. Sømme (1968) has suggested a subdivision of the relief into 3 intervals: a) < 20m, b) 20-50 m and c) ≥ 50 m. However, the data used falls into 4 natural groups so instead the result (Figure 7) was subdivided into 4 intervals: 0-25m, 25-50 m, 50-75 m and > 75 m, see Figure 7.

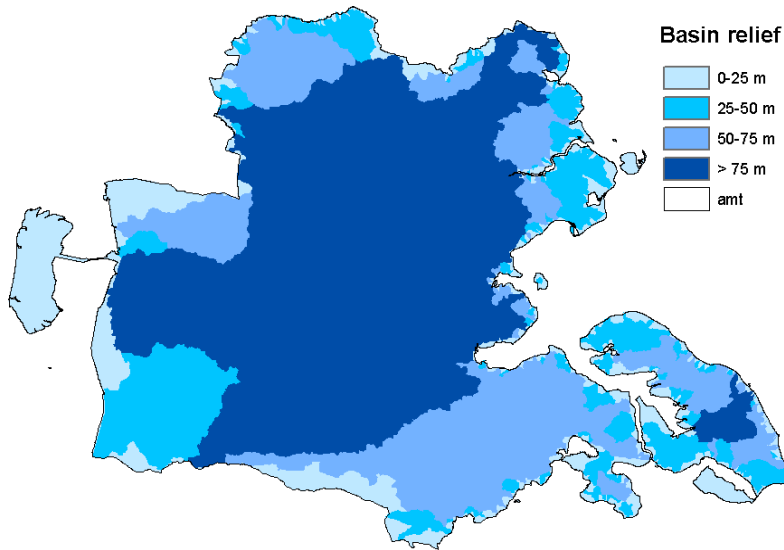


Figure 7 The basin relief within the County of Soenderjylland.

The slope of reach

The slope of reach for each river segment was derived in a two step procedure. First, the elevation value at each river junction was transferred from the DEM using functions from ESRI's 3D Analyst extension for ArcGIS (interpolate shape). Secondly, a Visual Basic script written for the occasion calculated the differences in elevation between the start point and end point for each reach divided by their euclidian distance. The result (Figure 8) was subdivided into 3 intervals according to natural data breaks: a) $< 0.5 \cdot 10^{-3}$ m/m, b) $0.5-1.0 \cdot 10^{-3}$ m/m and c) $\geq 1.0 \cdot 10^{-3}$ m/m.

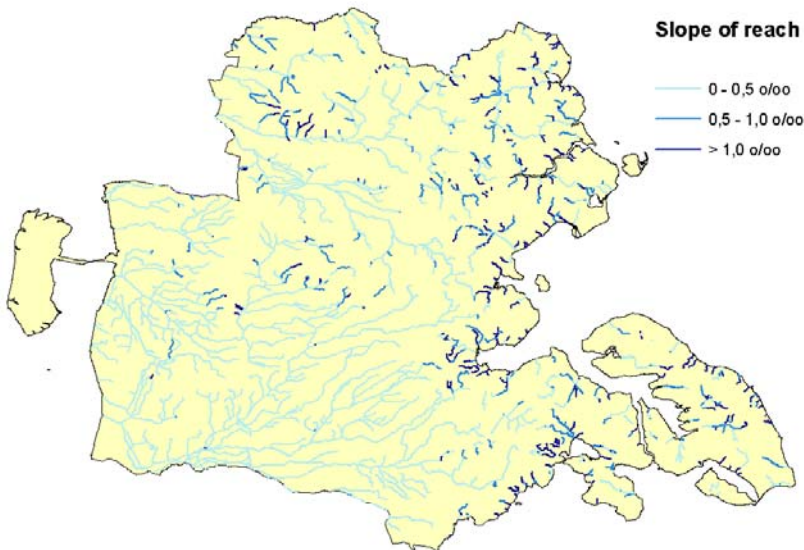


Figure 8 The slope of reach for each river segment within the County of Soenderjylland.

The sinuosity

The sinuosity was calculated for each reach as well from a calculation of the ‘curved’ length for each reach divided by the euclidian distance between the start point and end point. The sinuosity is subdivided into 3 categories as suggested by Leopold et al. (1964): a) straight, $SI < 1.05$, b) sinuous, $1.05 \leq SI \leq 1.5$ and c) meandering, $SI > 1.5$.

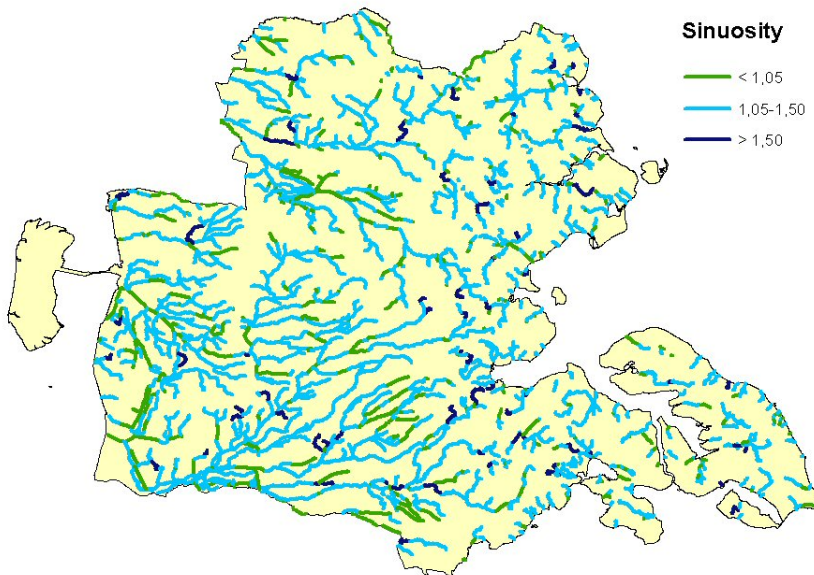


Figure 9 Sinuosity for each river segment within the County of Soenderjylland.

The pros and cons of the system

No doubt, the presented setup is a major step towards searches for specific physical conditions in the river system of Soenderjylland. However, during the creation of the database several essential problems related to data quality were observed. The problems may be grouped into the following categories:

1. Non-consistency of line data used to represent the locations and topology of rivers
2. Non-consistency of polygon data due to differences in scale

ad. 1 The line data used as reference data for the river segments is not up-to-date. They lack much in detail compared to the newest data from the National Survey and Cadastre on locations of watercourses at a scale of 1:10,000. Another problem is that the data set is not continuous. When topology is build several gaps are revealed within the river systems so due to this circumstance the river orders can not be calculated. Meanwhile, as long as the investigations so far do not focus upon flow of matters from one river segment and downstream this lack of continuity is not critical. But when modellings in this context become interesting a new corrected data set must be obtained. Although available such data are expensive to a county but will, hopefully, be available within a shorter period.

ad. 2 The inconsistency of polygon data can be visualized by simple overlay techniques. When displayed the imprecise overlaps are clearly visible. In relation to the data used in this project it is striking that thematic data collected by manual digitizing during the 1970s on lowlands, topsoils

and geomorphology (base maps at scales 1:20,000 to 1:100,000) can not sustain to be overlaid by thematic maps of newer origin. When polygons are not exactly overlapping and transferred as line events to the river theme they will create a very large number of very small events (in this case in total 15,262 events). Hereof approx. 1/5 have lengths < 20 metres and only 1/3 (5,654) have lengths > 100 metres! These numbers points directly to the big problem related to use several polygon themes of different origin and quality to establish a series of line events. The inconsistency of data is a matter to be taken very serious by the various data providers but discussions about which data set to use as a reference for other ones is delicate.

If the system should become operational it must, of course, be flexible to adjustments of interval values etc. The LRS is able to include point measurements as events along the rivers as well for operational monitoring of the rivers' hydrologic conditions.

Acknowledgement

We hereby thank the County of Soenderjylland for providing a line shapefile of the geometrical river network and associated attributes including a biologic assessment of the major river segments.

References

ESRI (2004): Linear Referencing for ArcGIS 9.0. Users Guide. ESRI, Redlands, CA, USA.

Miljø- og Energiministeriet (1998): Biologisk bedømmelse af vandløbskvalitet. vejledning fra Miljøstyrelsen 5/1998.

Hasholt, B. (1999): On classification and regionalisation of Danish watercourses. Dan. J. of Geography, Special Issue 1, 1999, pp. 81-87. Reitzel, Copenhagen.

Leopold, L.B., Wolman, M.G. & Miller, J.P. (1964): Fluvial Processes in Geomorphology. W.H. Freeman, San Francisco.

Mernild, S.H. & Hasholt, B. (2003): A test of classification and regionalisation of Danish watercourses. Dan. J. of Geography, 103(2), pp. 13-25. Reitzel, Copenhagen.

Sømme, A. (1968): A Geography of Norden. Oslo, Cappelen's Forlag.