

Review of Existing River Coding Systems

For River Basin Management and Reporting

European Coding Systems

WFD GIS Working Group

Version 1.9
Oct 2002

CONTENTS

Review of Existing River Coding Systems	2
1.1 Introduction	2
1.2 The US catchment coding system	2
1.3 The German catchment coding system (LAWA).....	5
1.4 The Norwegian catchment coding system (REGINE)	7
1.5 The Portuguese catchment coding system	7
1.6 The Irish catchment coding system	9
1.7 The Finish Coding System	11
1.8 Finish River network continua	17
1.9 The Spanish River Coding Systems	19
1.10 The Erica catchment coding system.....	22
1.11 Comparison of Key River & Catchment Coding Structures	25
1.12 Conclusions	28

1.1 Introduction

A number of member states have already adopted practices, most of which are somewhat similar, but which may be particularly suited to local geography. Not all existing coding systems are reviewed, just key ones and a sample of others to identify the issues. The objective is to derive a coding structure that is suitable to all member states. The Commission funded the Erica project for such a purpose. Hence it will be necessary to evaluate that projects findings. None of the existing used or recommended coding structures appear to be comprehensive in that they do not address each type of water body mentioned in the WFD. In addition they do not address pressures, status and impacts.

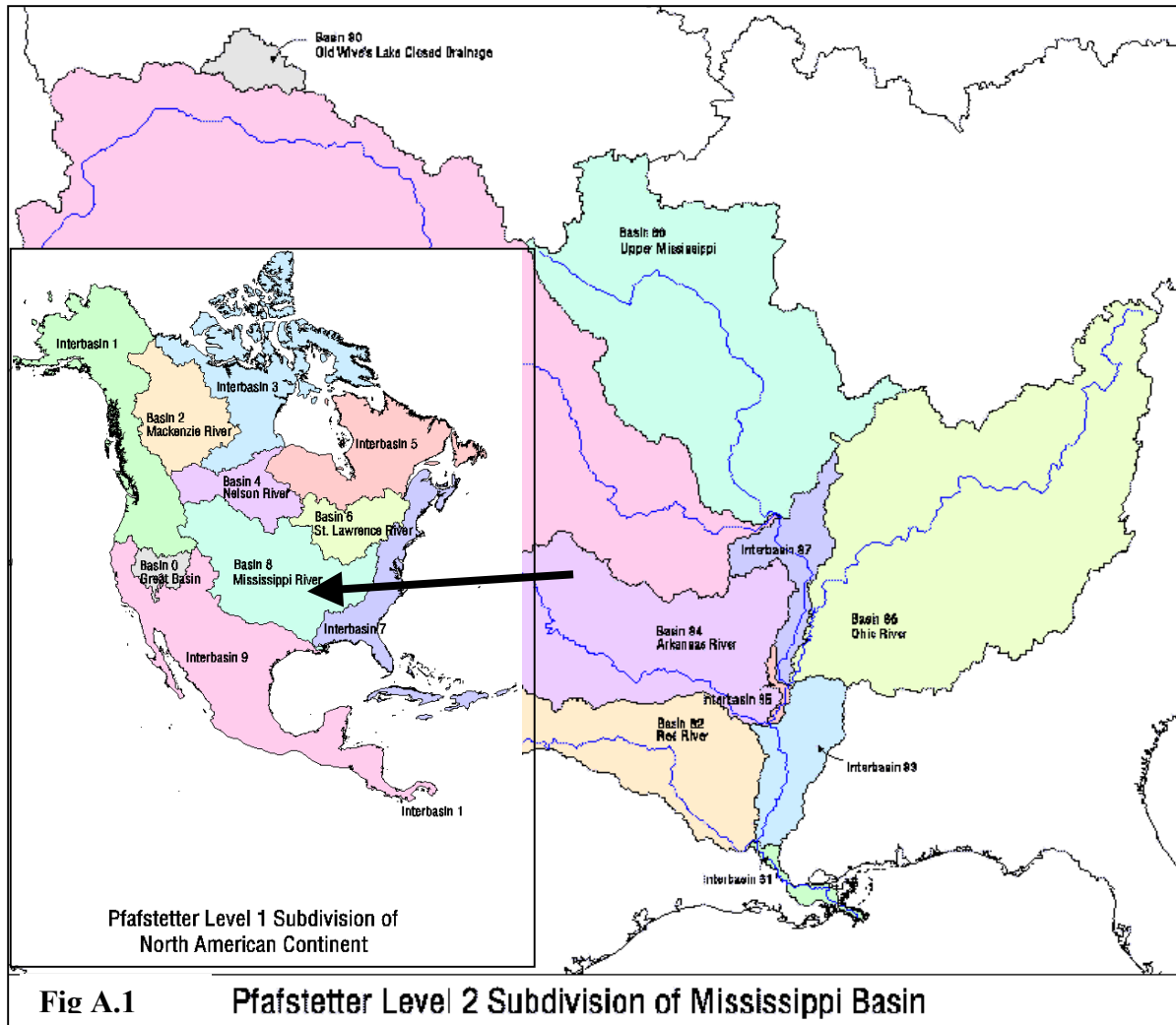
This section focuses on rivers and river catchments. These are more complex in structure than other water bodies due to the complexity of nested tributaries and nested catchments. Other water bodies will be dealt with later as will pressures and impacts.

A river (or river catchment) has many tributaries (or sub-catchments), which in turn have sub-tributaries (or sub-sub-catchments). If we start with Europe's largest rivers, this process can go down many levels. By using the same coding methodology at each level we can automatically determine the codes of the parent rivers or catchments for each tributary or sub-catchment. The number of levels to be used is largely determined by the size of the catchments required.

1.2 The US catchment coding system

The HYDRO1k data set was produced as part of a project involving the US Geological Survey's Data Centre and the United Nations Environment Information Programme. It is described by Kristine L. Verdin at <http://edcdaac.usgs.gov/gtopo30/hydro/P311.html> and synopsised here.

River basins and drainage networks were tagged according to the numbering scheme developed by Otto Pfafstetter, a Brazilian engineer. This system is based upon the topology of the drainage network and the size of the surface area drained. Its numbering scheme is self-replicating, making it possible to provide identification numbers to the level of the smallest sub basins. For a given location it is possible to automatically identify all upstream sub basins, all upstream river reaches, or all downstream reaches.



The area drained by a major river is subdivided into coded basins and inter-basins. The four largest tributaries, according to the criterion of area drained, form basins. These are assigned the numbers 2, 4, 6, and 8, in the order in which they are encountered as one goes upstream along the main stem.

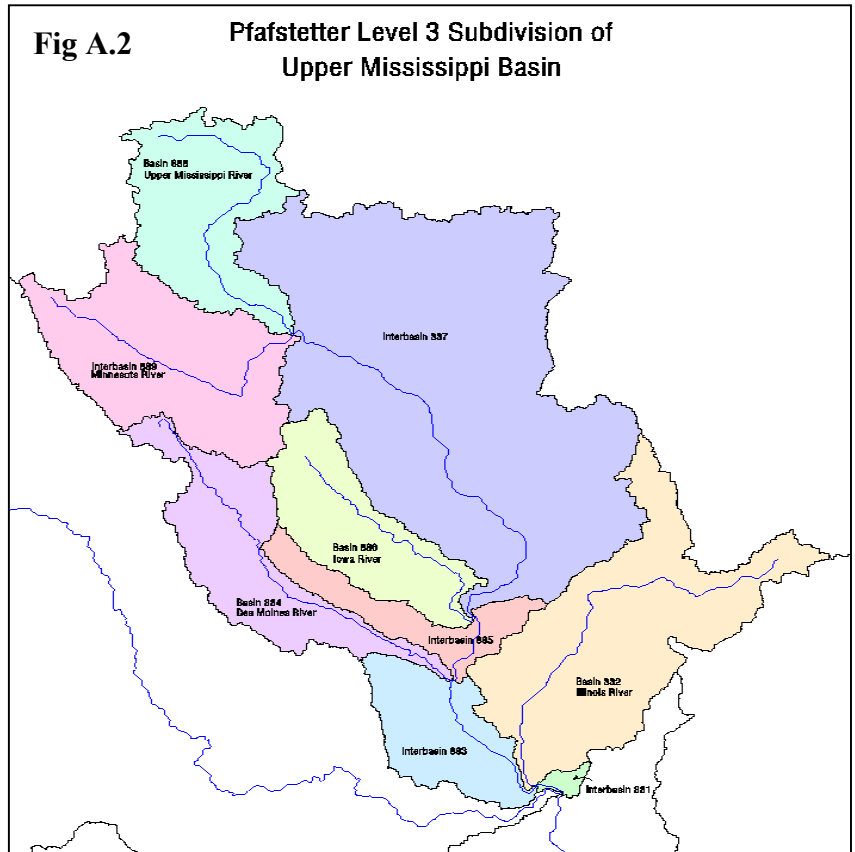
Next, the inter-basins are numbered 1, 3, 5, 7, and 9, again working upstream from the mouth of the main stem. Inter-basin 1 is the area drained by the main stem between the mouth of basin 2 and the mouth of the main stem. Inter-basin 3 is the area drained by the main stem between the mouths of basins 2 and 4. Inter-basin 9 always consists of the headwaters area of the main stem, and always drains more area than basin 8, by definition (9 being the remainder of the main stem and hence more significant than 8, a tributary). If a closed basin is encountered, it is assigned the number 0 (zero).

It should be noted that the river reaches along the main stem are identified by the inter-basin codes. Ultimately, the inter-basin codes define the full river network. They also identify the areas that drain in a diffused manner into that network

Figure A.1 demonstrates the subdivision of the Mississippi River basin into its first 10 sub basins. Note that the upper basin is a closed basin, otherwise there would only have been 9 sub basins.

Each basin and inter-basin is preceded, in the case of the Mississippi, by the digit eight (8). This digit designates the Level 1 Pfafstetter subdivision of the North American continent, (Pfafstetter, 1989).

A basin may be further subdivided by repeating the application of the same rules to the area within it. Thus, basin 88 of our example, the Upper Mississippi Basin, is subdivided into basins 882, 884, 886, and 888 the Illinois, Des Moines, Iowa and Upper Mississippi Rivers), and inter-basins 881, 883, 885, 887, and 889. This is shown in Figure A.2.

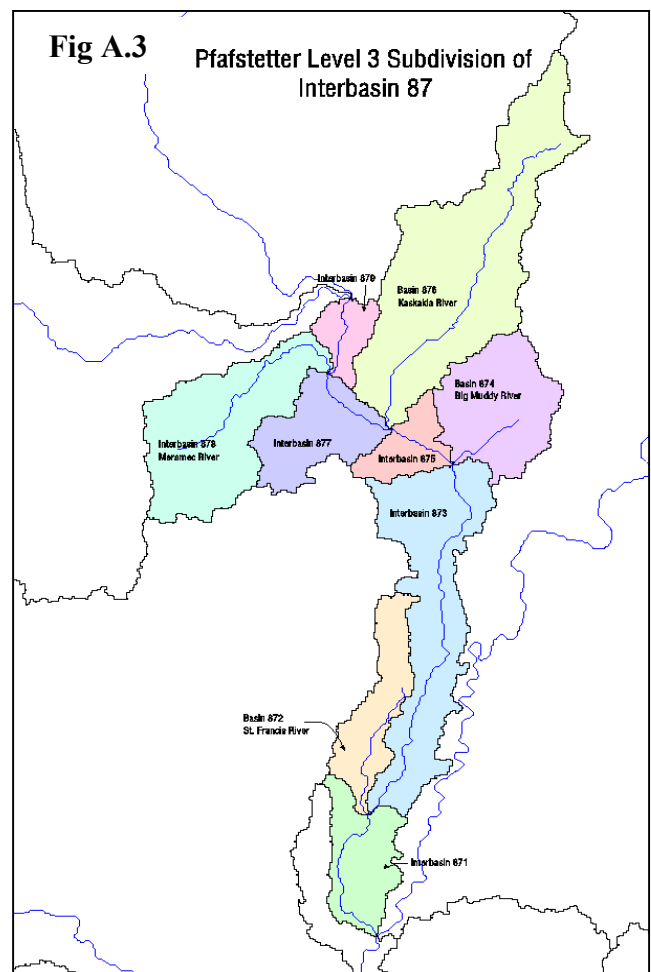


An inter-basin, see Figure A.3, is subdivided by identifying the 4 tributaries within it having the greatest drainage areas. They are numbered from downstream to upstream, regardless of whether they enter the main stem on the right or left bank. Thus, inter-basin 87 contains basins 872, 874, 876, and 878. As in the case of the main stem, the areas drained by the intervening reaches are numbered as inter-basins 871, 873, 875, 877, and 879.

further subdivided until four tributaries can no longer be found.

Subdivision of coastal inter-basins requires examination of the four largest streams flowing to the coast. These are numbered in a clockwise order as 2, 4, 6, and 8 and the appropriate basins defined.

Sub basins down to Level 5 Pfafstetter units were extracted. This produced 5020 Level 5 Pfafstetter units for North America with an average surface area of 3,640 sq. km. The resolution of the DEM did not merit finer extraction of basins.



Upstream-downstream dependency between locations can be inferred by examining identification numbers. Consider a cannery situated on a river in inter-basin 8873. See Figure A.4. Will a new dam at a location in 8885 affect flows to the cannery? The identification numbers reveal that the dam site is upstream of the cannery, without need to reference a map. The match of the leading digits, 88, tells us that the two locations are in the same basin. Beyond the matching digits, 85 is greater than 73, and therefore the dam lies upstream of the cannery.

Will the dam affect the irrigation diversion of a farmer at 8834? No, it will not, even though we see that 85 is greater than 34. This is because the trailing 4 indicates a tributary off the main stem and above any flows influenced by the dam. Thus, river reaches affected by the dam will have a match of leading digits, 88, and trailing odd digits less than 85.

It can be seen from these examples that simple rules to check digits with tests of "odd" or "even", and "less than" or "greater than", can quickly isolate areas of interest for a particular investigation.

1.3 The German catchment coding system (LAWA)

The German method of catchment coding has been established by the Länderarbeitsgemeinschaft Wasser (LAWA). The coding system is hierarchical and purely numeric. The system is, as far as possible, related to hydrological catchments. Every subdivision of a catchment into sub-catchments introduces a new digit. The system is very similar to the Pfafstetter system, except that the numbering is carried out in the other direction. Thus the sub-basins are numbered from the headwaters to the river mouth.

Figure A.5 illustrates the coding of the principal drainage systems in Germany. The numbering starts at the Donau (code

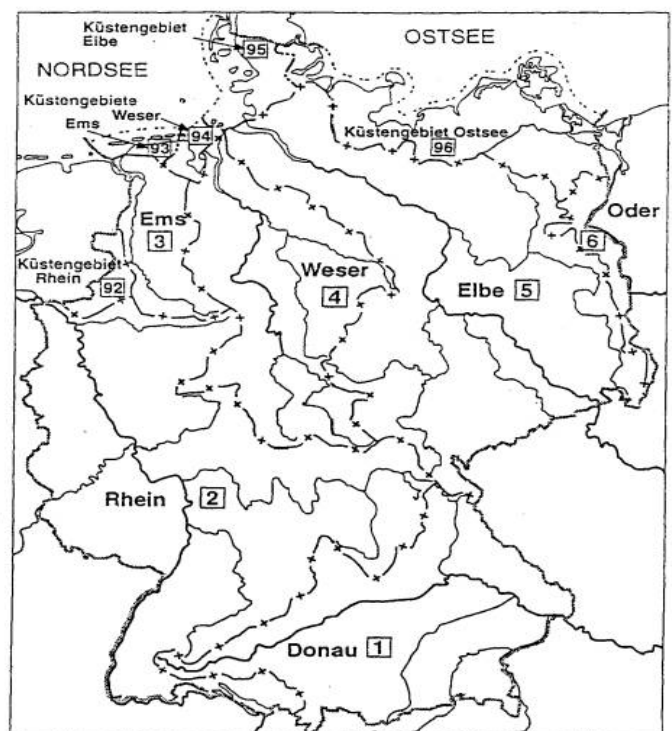
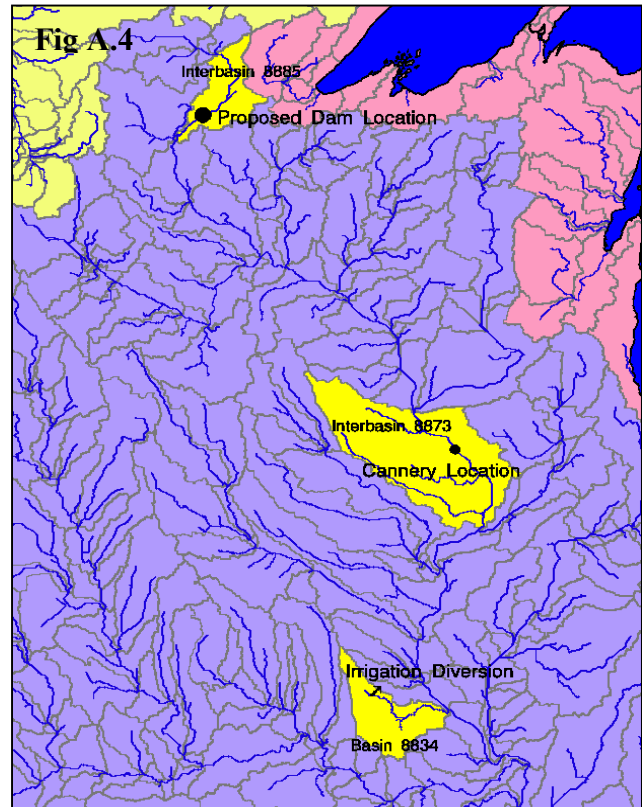


Fig A.5 Coding principal German drainage systems. From: Länderarbeitsgemeinschaft Wasser (1993)

1), and ends at the Oder (code 6). Every catchment can be subdivided into 9 sub-catchments. This is shown in Figure A.6 by an example of the River Rhine catchment (code 2). Sub-catchment no. 1 (code 21) characterises the headwater area. The even numbers 2, 4, 6 and 8 stand for tributary catchments. The odd numbers 3, 5 and 7 express inter-catchment areas along the main river. These inter-catchment areas are hydrological catchments, since all water drains to one outlet point. Number 9 represents the area where the river flows into another river or into the sea.

The subdivision of inter-catchment areas is made following exactly the same methodology. Figure A.7 shows the subdivision of the inter-catchment area 27. Again, even numbers stand for tributary catchments (codes 272, 274, 276, 278) and odd numbers for inter-catchment areas (codes 271, 273, 275, 277). Catchment 279 is the area near the outlet of inter-catchment area 27.

Between the outlets of the principal drainage systems, there are coastal areas. These do not belong to any principal catchments. They are drained directly into the sea by smaller rivers. As Figure A.5 shows, all of these areas receive the digit 9 on the highest level of hierarchy of the LAWA system. On the second hierarchical level, these areas are subdivided according to geographical neighbourhood, e. g. the coastal area between Ems (code 3) and Weser (code 4) is subdivided into area 93, neighboured to the Ems, and area 94, neighboured to the Weser. The borderline between these sub-areas is subjectively determined.

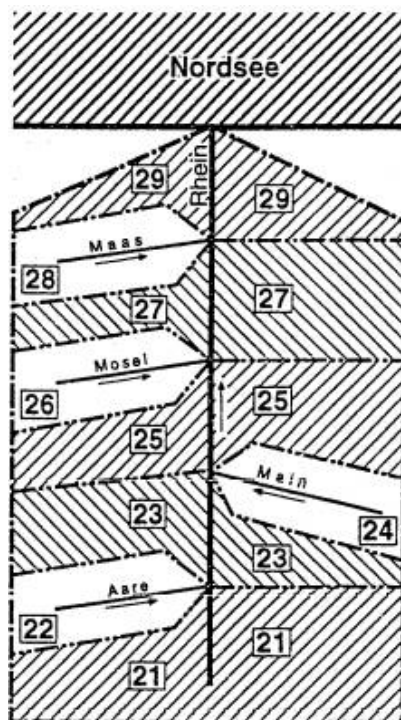


Figure A.6 Coding within a principal catchment (the Rhine).

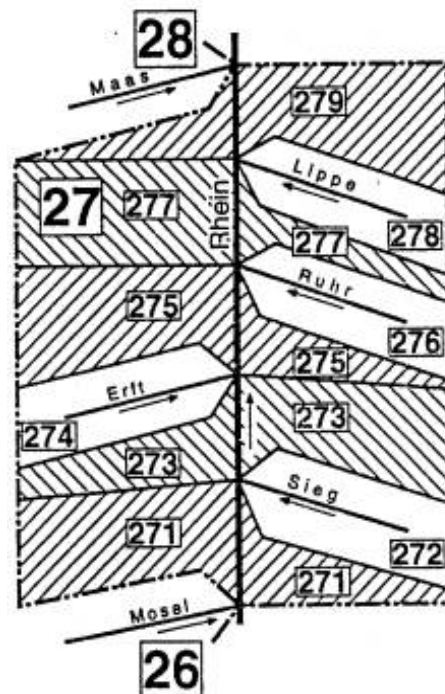


Figure A.7 Coding within an inter-catchment area of catchment 27.

1.4 The Norwegian catchment coding system (REGINE)

The methodology of the Norwegian Register of Catchment Areas (REGINE) coding system is shown in Figure A.8. As a first step, the whole territory of Norway is subdivided into hydrological water system areas (e. g. 016 in Figure A.8), which are not identical to hydrological catchments.

The water system areas on the highest hierarchy level are subdivided into two parts, a main catchment with one single outlet (016.Z in Figure A.8) and a remaining area drained by diffuse drainage systems or minor watercourses along a lake or the coastline (016.0 in Figure A.8). Thus in this step a qualitative distinction is made by choosing either a letter or a number. One digit/character is added.

The main catchment areas can be subdivided into 24 sub-areas at most, using the letters A to Y. 016.Z in Figure A.8 is subdivided into 016.A to 016.N. The areas of diffuse drainage can be subdivided into 9 sub-areas, employing the numbers 1 to 9 (016.1 to 016.4 in Figure A.8). In both cases the subdividing is done without addition of a new digit.

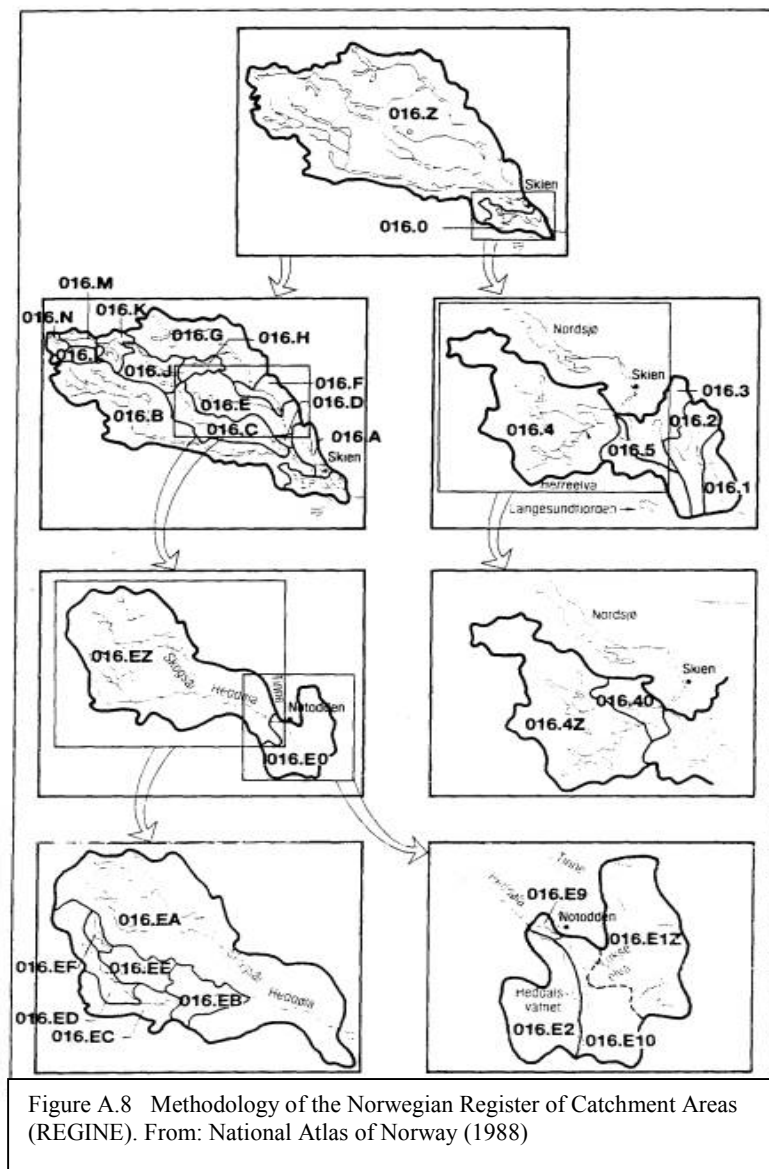


Figure A.8 Methodology of the Norwegian Register of Catchment Areas (REGINE). From: National Atlas of Norway (1988)

The same methodology, switching between a qualitative distinction (and addition of a new digit) and division into sub-areas (without addition of a digit) is now applied for the lower levels of hierarchy. Area 016.E in Figure A.8 is subdivided into a main catchment, 016.EZ, and a remaining area, 016.E0. Catchment 016.EZ can be subdivided into sub-area 016.EA to 016.EY, and unit 016.E0 can be subdivided into 0.16.E1 to 016.E9, and so on.

1.5 The Portuguese catchment coding system

Before embarking on the use of this coding system, other systems were examined by Portugal. It was decided that both the Pfafstetter and the German LAWA system were

DRAFT - Prepared by P. Britton 04/11/02

As part of EU GIS Working Group Guidance Document
adequate for implementation in Portugal but that the LAWA system would require
modification in the treatment of coastal areas. Hence the Pfafstetter system was adopted.

The ERICA system was also reviewed by Portugal. It was deemed to be being too difficult to identify up to 49 main tributaries for each of the European rivers. The Portuguese evaluation of ERICA determined that 49 main tributaries have to be identified, even for smaller rivers. However, it is possible to interpret the ERICA documentation such that a lower number of tributaries can be used.

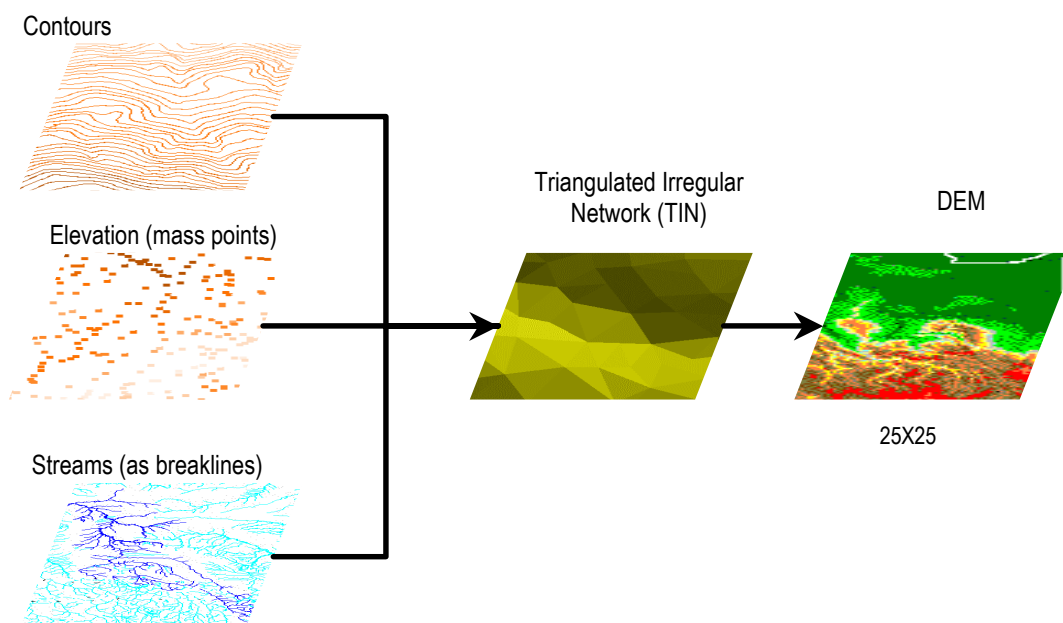
The system used by Portugal is similar to any of the above systems but in particular to the Pfafstetter system, which it follows with the following exceptions.

1. The '0' digit is not used to code closed basins, but to identify the level beyond which coding is no longer possible with existing topological information. However, this variation from Pfafstetter was more a software scripting procedure than a coding requirement.
2. A north to south rule was use when initially numbering the Portuguese major river basins and intermediate coastal regions.

Closed basins were filled to produce a continued flow beyond the depression. That is, a spill point could always be found from which a continued flow was reasonably generated.

Using this system, all of the rivers in Portugal have been successfully identified for catchments that exceed 10km², and smaller where mapping allowed.

This work was based on several digital elevation models, which were generated using contours (with height increment between contours of 10m), spot heights and streams as break-lines. The methodology and data used were:



The source data (scale 1:25000) used to create the DEM's is produced by the Portuguese Army Geographic Institute (<http://www.igeoe.pt>). The height increment between contours is 10 meters and the height points have a precision of 1 meter. The resulting DEM has a vertical resolution of 2 meters.

For a better resulting TIN the “weed tolerance” was varied when working in flat areas or not (to avoid flat triangles).

Generally between 3 and 7 levels of Pfafstetter coding was required. However, in some instances, 9 levels were required to identify all of the rivers visible on 1:25,000 scale mapping.

Maximum area of a catchment: 83611875.000 m²

Minimum area of a catchment: 3750 m²

Number of catchments identified in Portugal: about 35500

Where 4 tributaries could no longer be found on 1:25,000 maps, then the coding terminated. However, this always produced catchments that were less than 10 km².

1.6 The Irish catchment coding system

One coding system used in Ireland is for the purpose of river hydraulics and physical maintenance. The main channel is designated C1. Tributaries are designated by their distance from the outfall along the channel into which they flow, e.g. C1/21/17/5 is a third order tributary. C1/21 is the 21st tributary of the main channel from its outfall. C1/21/17 is the 17th tributary of this tributary etc. All channels are marked from the outfall at 100 yard (91 meter) intervals. It is possible to identify a reach as C1/21/17/5 from chainage 21/00 to 24/35 i.e. between 2100 yards and 2435 yards from its outfall. This system provides a lot of topological information. This system has been used for many decades. To date however, it has not been incorporated into a GIS environment.

In the 1970's a river coding system was established for river quality monitoring. It is based on hydrometric areas, and like the Norwegian system, these are not hydrological catchments, they can consist of a number of independent catchments. There are 40 hydrometric areas, which are numbered in a clockwise manner around the coastline, including Northern Ireland. Within each hydrometric area, rivers and major tributaries are identified by the use of one character and two numerics. The character is based on the first letter of the river name and the numerics are used to separate rivers within that hydrological area that have a common first character in their name.

Thus for example the River Suir, which is in Hydrometric Area 16, has a code of 16-S02. The River Shanbally, also in Hydrometric Area 16, has a code of 16-S01. To meet with the requirements of the Water Framework Directive, this coding structure may need to be extended to deal with an increase in the number of river catchments to be identified.

This coding system is used primarily to identify monitoring station along the river network. These are assigned using initial intervals of 100, to allow further interspersed stations. Thus a typical station would be referenced as 16-S02-0200, and another as 16-S02-0300. The codes carry no topological information about the rivers or their connectivity. Such information is maintained only within a GIS environment. However, the monitoring data can be readily integrated with GIS due to the well established standard river monitoring station codes.

All topological information about river reaches and their connectivity is held in Catchment Envisage GIS. Catchment Envisage software extends the coding system by introducing codes for each river stretch. Stretches are automatically numbered in intervals of 10 to allow future sub-division. The numbering process works downstream along the main channel. Thus the stretch from the river source to its first confluence is identified by adding the 3 digits 010, the next stretch by adding 020, etc. Thus a river stretch may be identified as 16-S02-020. Tributaries, which have not received independent monitoring codes, from the EPA, are assigned the code of the river into which they drain. The stretch numbering for these is then a continuation of the sequential downstream numbering that was applied to the main river. Rivers are classified as being either 'River', 'Tributary' or 'Tertiary'.

This coding determines national stretches. It is based on EPA designated codes and on river topology. Catchment Envisage allows local sub-division of these codes by using a further two digits. Hence, if a small stream associated with discharges is to be recorded, the national stretch can be subdivided. Local stretches are automatically numbered using intervals of 10 to allow further sub-divisions. For example, if national stretch 16-S02-020 is to be divided by a locally identified stream, then the upstream local stretch would be identified as 16-S02-020.10 and the downstream stretch by 16-S02-020.20. Within the GIS, both national and local stretches are maintained graphically, making it possible to report against either.

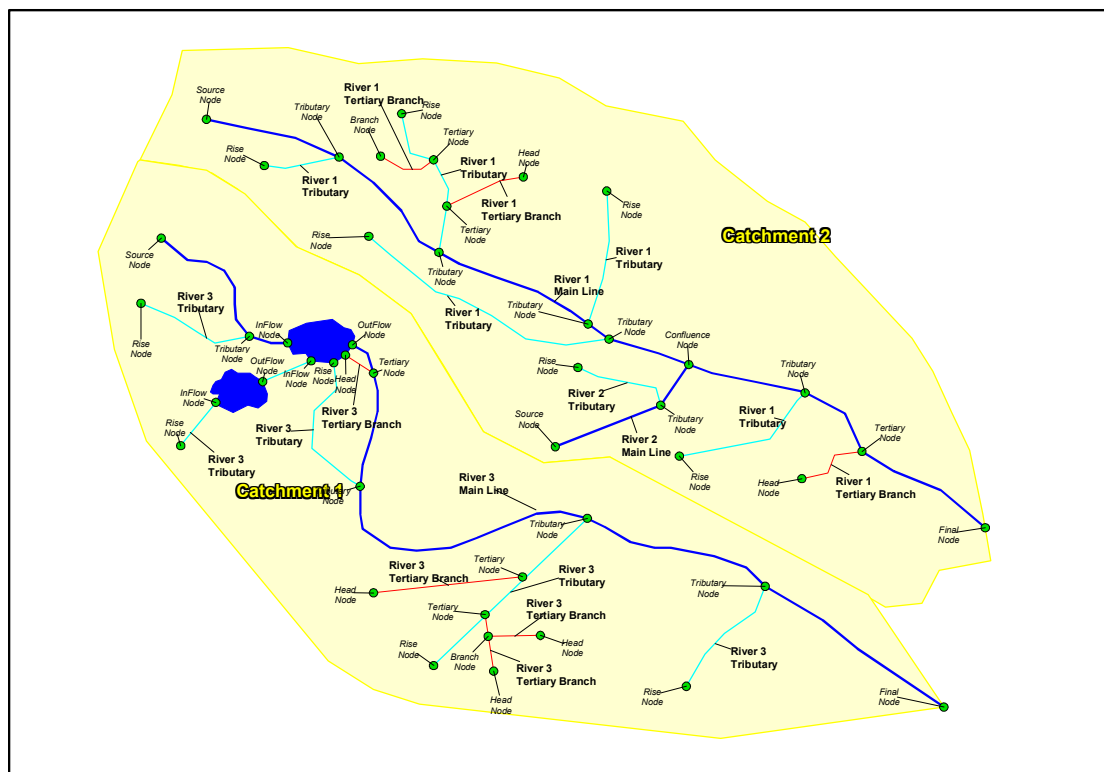


Fig.A.9 Catchment Envisage Typical River, Catchment & Node Layouts

Catchment Envisage also has a systematic numbering system for each node at confluences, sources, outfalls, etc. Node coding is not discussed here as the main environmental focus is on river stretches.

River topology is managed directly within GIS, the structured coding is more to assist in the assignment of unique codes. Through the GIS, it is possible to determine drainage routes and to determine if features such as discharges will impact on other downstream features. This is the system used primarily for catchment management. Figure A.9 identifies the type of topology that is maintained within Catchment Envisage GIS.

1.7 The Finish Coding System

The Finish codes start with the identification of regions (basins) and then deal with the identification of networks (river and stream) within these regions. A final component of the code gives us an indication of the size of the upstream catchment. The initial component of the code (basins & sub-basins) is also used to start lake coding.

At the highest level, the 74 main river basins, larger than 200 km², are identified and numbered. These main basins are then divided into sub-basins down to three levels of hierarchy. Also, 177 small coastal basin boundaries were determined, but without further division into sub-basins. This basin boundary inventory was based on 1:50 000 topographical maps. The boundaries were delineated manually, then digitized..

The river basins are divided into sub-basins and basin parts, beginning at the lower course and following the channel, which is selected to be the main channel; then the tributaries are numbered in a clockwise manner. The longest channel is usually selected as the main channel. The second level of the hierarchy refers to tributaries, and the third to creeks and minor rivers.

River basins codes form the basis for codes in several other national databases (in the Finnish Environment Institute). The river basin database was completed in 1991. Each of the catchment parts has a unique code by which it can be identified. In a later phase, the sea areas were added to the database

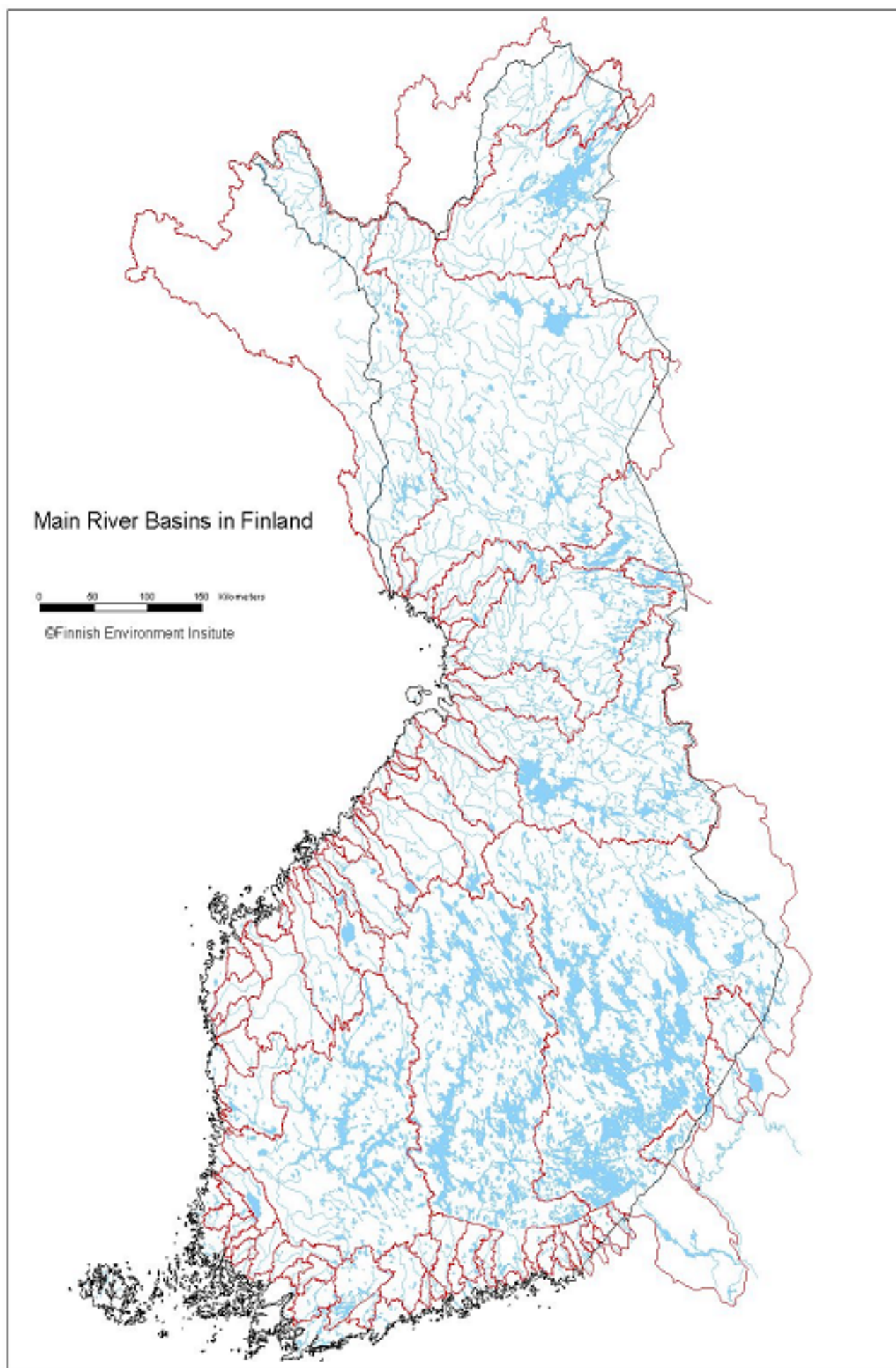
A pilot project on a water channel register was started in the spring 2000 and concluded in the autumn 2001. This was aimed at producing a list of rivers with catchment areas greater than 10km². This was based on extending the river basin codes with codes that focused on the topology of the contained river networks. For this purpose, the watersheds were modelled using the shoreline from the 1:10,000 terrain database and a height model with an absolute accuracy of about 7m.

Below third level sub-basins, river networks are divided at junction points into separate segments, each of which receives its own code. The river code (e.g. 42.012-010201.00.2) consists of an 18 character combination formed by the catchment area code, river network continuum hierarchy and the size class of the catchment area.

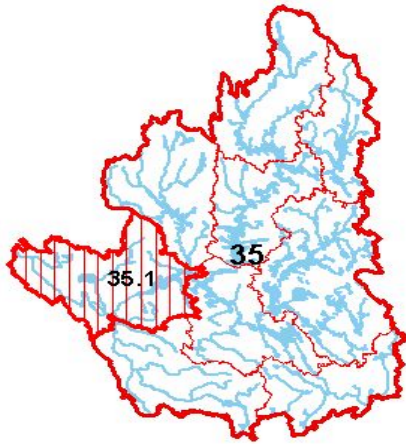
E.g. **42.012**-010201.00.2 = River basin code of the 3rd sub-division
 42.012-**010**201.00.2 = First side river (hierarchy level II)
 42.012-01**020**1.00.2 = Second side branch of a side river (hierarchy level III)
 42.012-0102**01**.00.2 = First tributary of a side branch (hierarchy level IV)
 42.012-010201.**00**.2 = Serial number (segments, upper course, bifurcations...)
 42.012-010201.00.**2** = Size class (and/or river type) of the catchment area

Because the most common hierarchy models, e.g. Strahler's, begin from the upper course of river network downwards, when new rivers are added to a river network system, the

As part of EU GIS Working Group Guidance Document river codes must almost always be changed at the same time the hierarchy model is readapted. With a specific hierarchy model created for a river code, main and side water rivers can be reasonably permanently classified and river network continua can be bound to each other. In addition, the river code contains an informative character describing the significance class of the rivers, i.e. the size of the upper catchment area. The catchment area size is classified into six classes: (1) $<10 \text{ km}^2$, (2) $10\text{-}50 \text{ km}^2$, (3) $50\text{-}100 \text{ km}^2$, (4) $100\text{-}1000 \text{ km}^2$, (5) $1000\text{-}10,000 \text{ km}^2$, (6) $10,000 \text{ km}^2+$, (0) not known.



The River Basin hierarchy



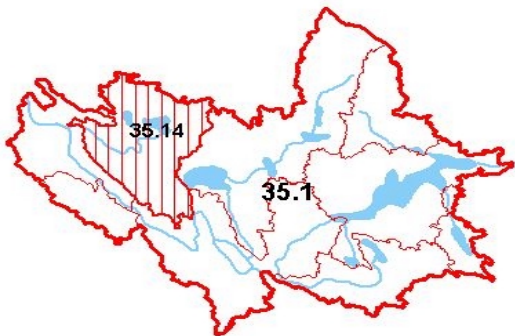
Main river basin

code name

35 Kokemäen joki

total area 27 046 km²

lake percentage 11 %



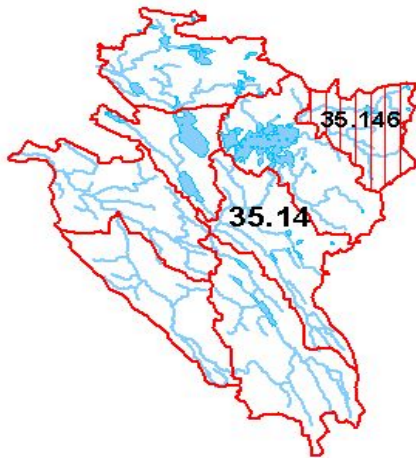
Sub-basin 1st level

Code name

35.1 Kokemäen joen alue

total area 3 679 km²

lake percentage 5,3 %



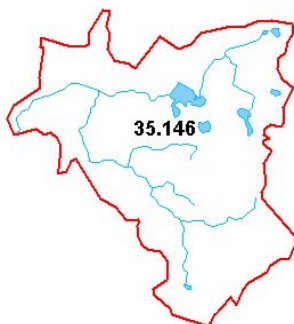
Sub-basin 2nd level

Code name

35.14 Harjunpäänjoen osa-alue

total area 506 km²

lake percentage 4,7 %



Sub-basin 3rd level

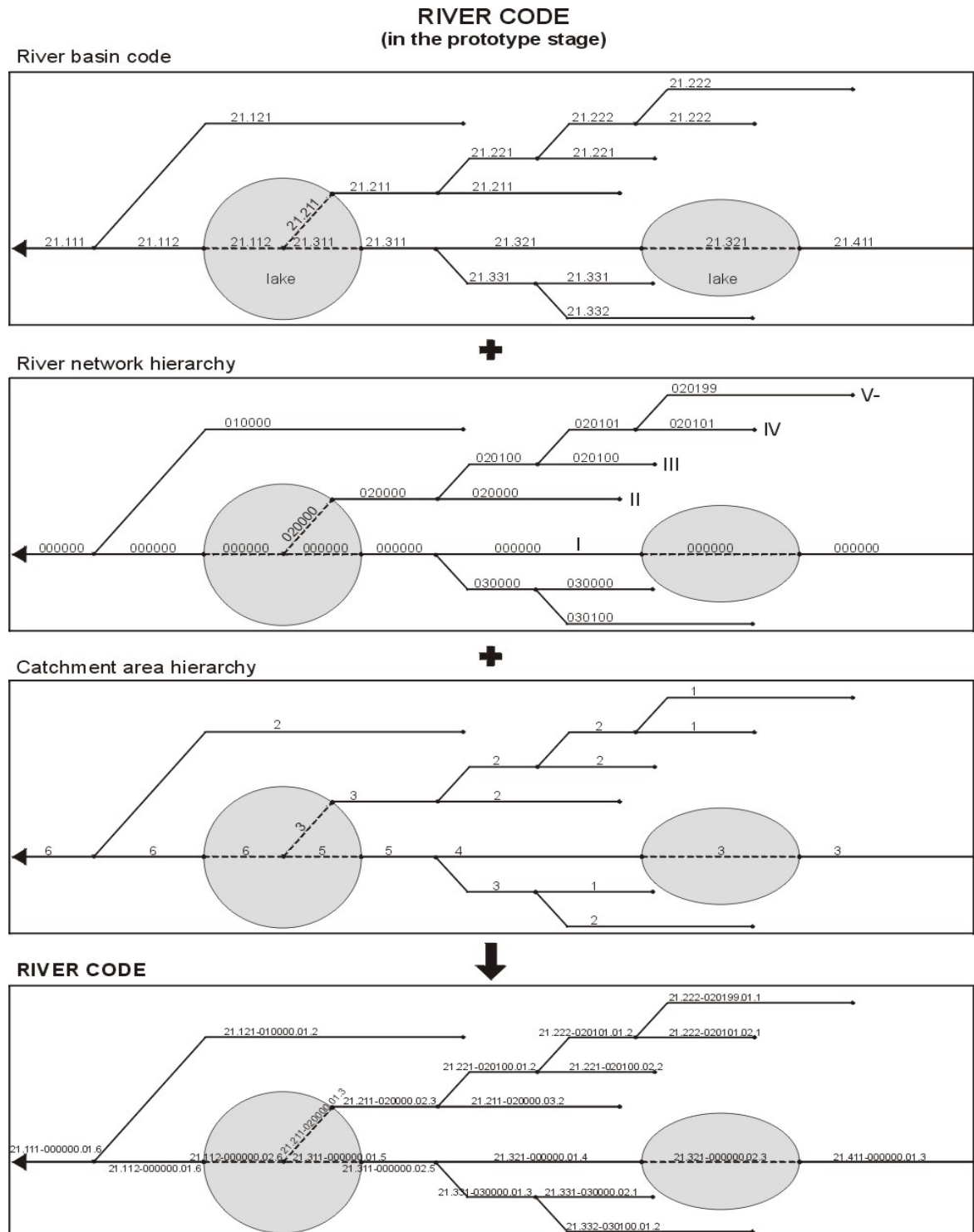
code name

35.146 Tyvijoen osa-alue

The river code systematics, based on a five-step river network continuum hierarchy, is built on the main rivers (min. catchment area 50 km²). The main river (000000) is usually

As part of EU GIS Working Group Guidance Document the most notable and longest water channel. The hierarchy code section is formed, on each level, with two characters that make it possible to have 100 rivers (00-99) on each level.

The catchment area codes and the river network continuum make it possible to locate a river topologically, regionally and nationally. The river size of the upper catchment area can be used as additional selection criteria. The proposed river coding makes it possible to expand the river network register, if necessary, with new rivers without a need to update existing river codes. A river network subsequently added receives the first free available code for the water channel system, catchment area and hierarchy level.



Lake Coding

There are over 200 000 lakes in Finland and over 59 000 of these are larger than 1 hectare. The fundamental data of the lake database was collected during the years 1986-1991 at the Hydrological Office of the National Board of Waters and the Environment (later Finnish Environment Institute). All lakes larger than 1 hectare were recorded in the register. Each received a code, based on the river basin code. Also, the coordinates of the visually defined center point of the lake were added to the database.

The lake code is of the format AA.BBB.C.DDD. It is constructed so that the first part AA.BBB is the third hierarchy level river basin code, while DDD is a number that identifies this lake among the lakes of this river basin part, for example 34.03.1.001 Pyhäjärvi is located in the river basin, which code is 34.03 and it is natural lake (the code 1). Some lakes belong to two or more river basins. The list of lakes (lake database) includes also the codes and areas of these parts of lakes.

1.8 Finish River network continua

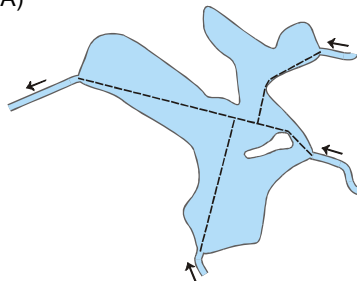
Though not strictly relating to coding, this section deals with the valuable experiences in Finland when establishing the river network continuum to be coded.

To make network analyses possible, a topologically uniform, one-dimensional river network in vector format must be generated from the rivers as well as the lakes. River polygons are converted into one-pixel wide raster lines. The water channel continua, flowing into and out of lakes, are modelled according to the principle of shortest distance by using cost-surface functions. If necessary, the main water channel that crosses the lake is digitised with the help of main flow directions, deeps and other local knowledge using the principle of shortest distance (A). This alternative seems the most practical, although local adjustments will probably have to be made in any case.

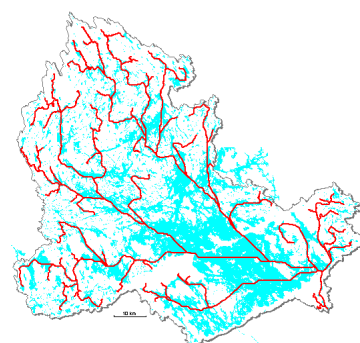
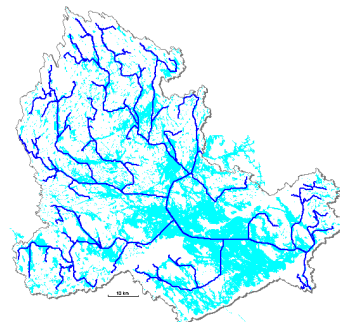
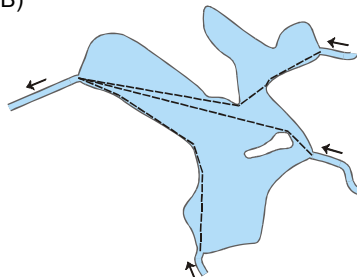
Another alternative (B) would be to define river network continua according to the shortest route for each directly to the discharge point of the lake. However, this alternative does not provide a “truthful” picture especially in large lake basins with numerous islands. Alternative B also distorts the total length of a water channel continuum network to be longer than in “real life”. On the other hand, the cumulative length of an individual water channel continuum branch is shorter.

RIVER NETWORK MODELING IN LAKE AREAS

A)

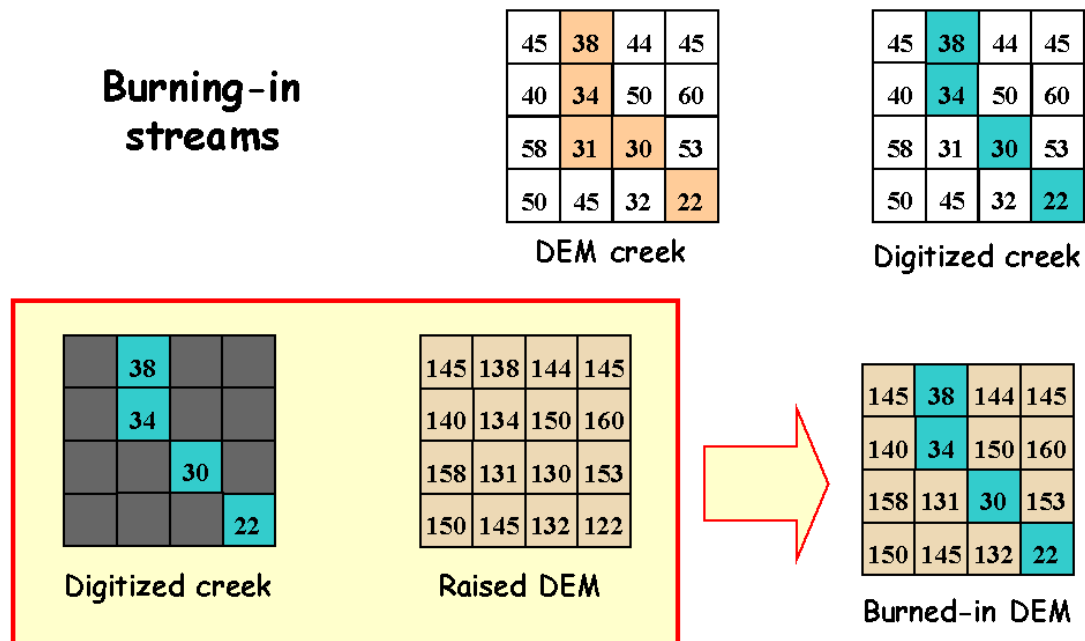


B)



Incorporating streams

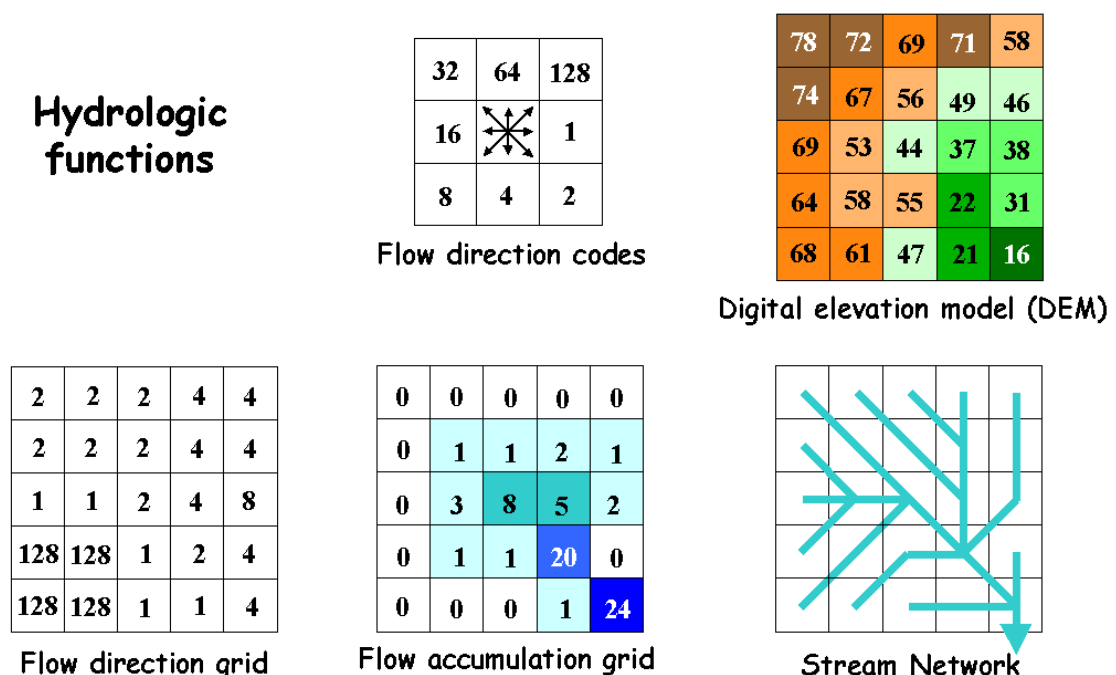
River network continua converted into raster form are incorporated into the height model with “depth gradation” according to the stream width. Wider rivers are incorporated deeper than smaller ones so that the depth gradation follows the depth hierarchy of the water channels. In practice, the widest rivers are left at original heights while the smaller water channels and the land pixels surrounding them are raised. In this way, the location accuracy of the flow and catchment area analyses made with the height models improves significantly.



Flow and catchment area algorithms

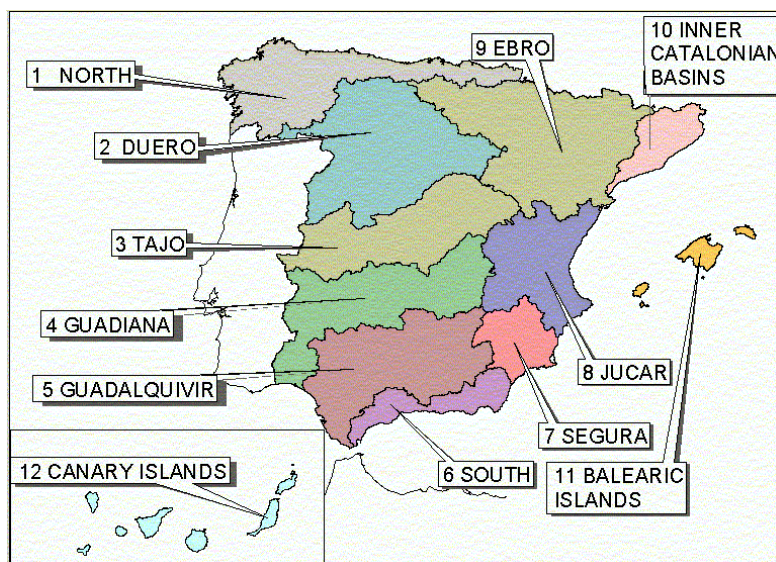
Water channels meeting the 10 km² catchment area criterion of the river network register are defined from the height model incorporated into the shoreline material of the terrain database by using special flow and catchment area functions. As a result, a river network formed by pixels having a catchment area of 10 km² is produced. However, it still also contains additional pseudo water channels of lake basins.

Hydrologic functions



1.9 The Spanish River Coding Systems

Spain currently uses two systems for river coding, a classical one, based on topological and hydrological criteria, and a more recent one, based on the rivers length. The first was created by CEDEX in the sixties; the second was created more recently by the Environment Ministry for water quality monitoring.



1. The rivers' decimal classification CDR (CEDEX).

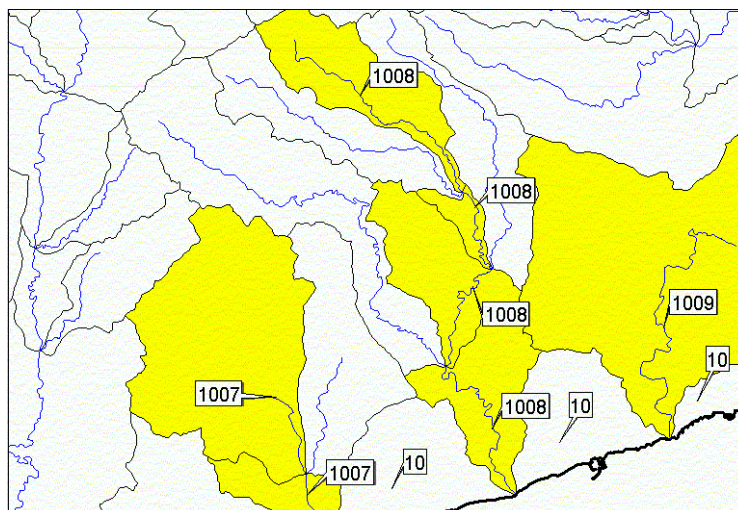
This river coding system was obtained manually in 1965 from 1:50.000 topographical maps. This was reviewed, updated and digitised in 1996. In each watershed one, and only one river, was coded with the same code. The river basin district defines the first element of the code. There are 12 main river basin districts, originating from the beginning of the 20th century.

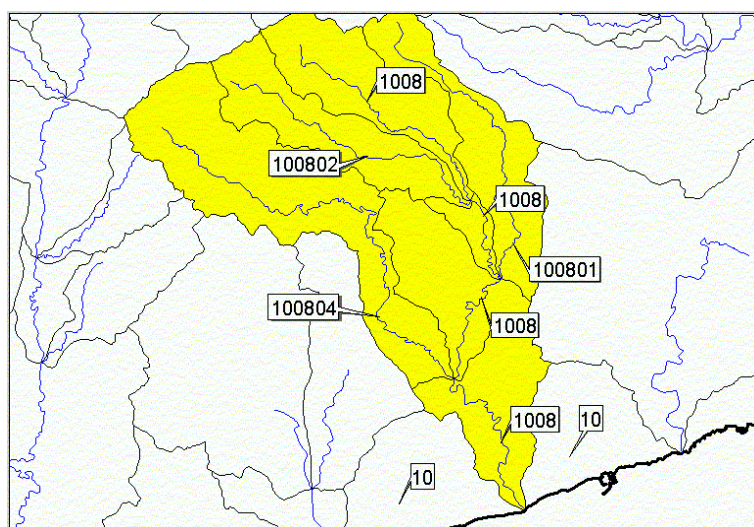
a) Cdr1: decimal river classification 1.

The code is generated by pairs of numbers in a nested system. Within each river basin district, rivers that drain to the sea are first level rivers and have two pairs of numbers, turning round the Spanish coast. The inter-basins between this first level rivers are not yet uniquely encoded and receive the code of the river district. The same occurs with closed basins.

Here is an example in river basin district 10:

Tributaries are numbered as second level rivers, adding a new pair of numbers, odd for left hand tributaries and even for right hand ones, moving from source to exit.

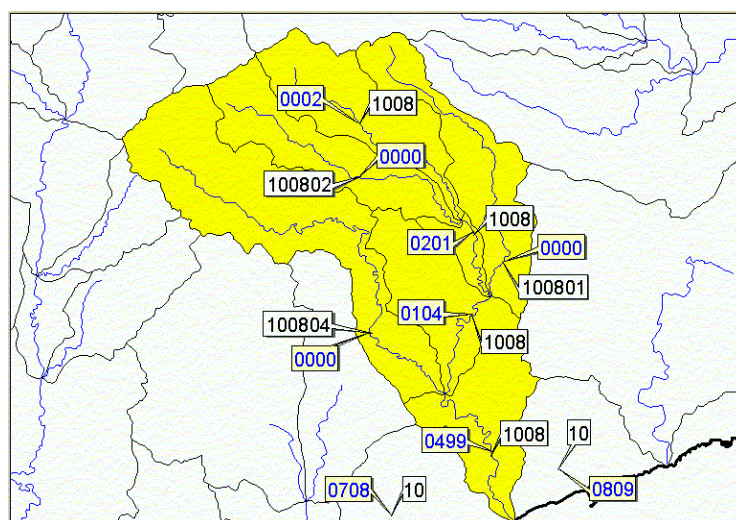




Tributaries of tributaries receive new pairs of numbers following the same rule. Currently coded basins for the whole Peninsula reach six levels and cover basins with up to 50 km² surface.

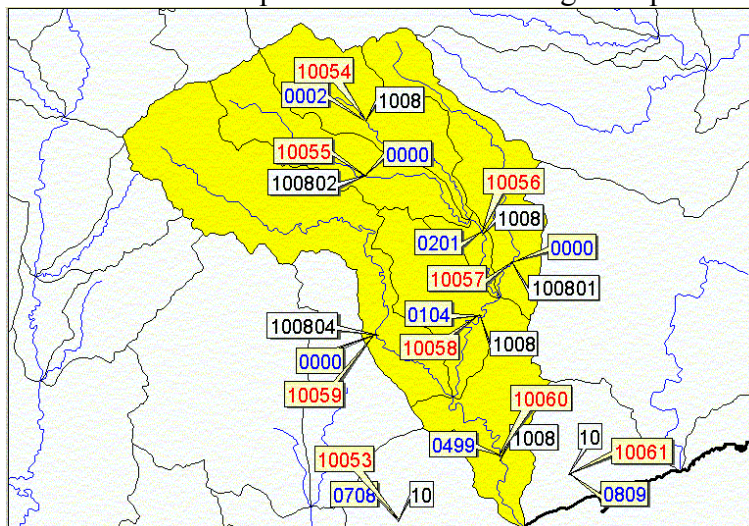
b) Decimal river classification 2: cdr2.

This number codes segment of rivers between tributaries. It always has 2 pairs of numbers, corresponding to the tributaries that define them. For example, the segment of river cdr1 = 1008 defined by tributaries cdr1 = 100801 and cdr1 = 100804 is cdr2 = 0104. If the river segment is delimited by the source, the pair of numbers is 00, if it is limited by the exit in another river, 99. Inter-basins receive the numbers of the first level rivers that define them.



c) number of the catchment:

This third code is the last one, and numbers each basin and inter-basin from upstream to downstream all around the basin district, following the order of the first level rivers. In fact it is normally used as key code in databases.



This coding system has been applied to all the catchments in the Spanish peninsula bigger than 50 km².

Advantages and disadvantages of the decimal classification coding:

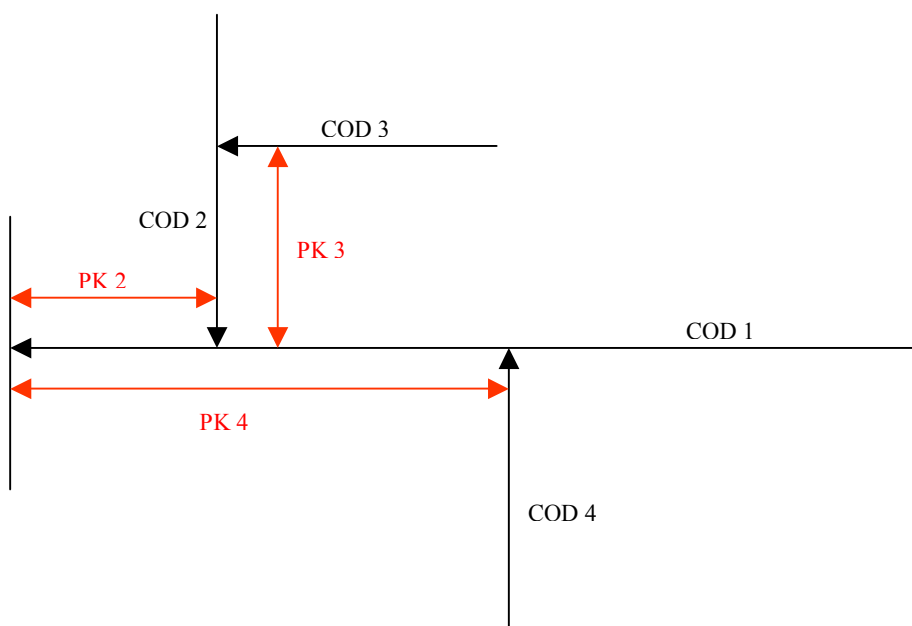
An advantage of this coding system is that the topology of the rivers is perfectly controlled (upstream, downstream, left hand, right hand, nesting possibilities, etc.). Another important feature is that the catchments were defined following hydrological criteria. In Spain water quantity is very important, and lots of big catchments produce very little flow. Often small basins have tributaries with bigger surface area, but smaller discharge values; this issue is reflected in the names of rivers. These problems are taken into account in this coding system.

A disadvantage is that to add a new basin, you have to move a lot of codes in the river district. For water quantity monitoring, almost any important river is included, but for water quality monitoring discharges from small basins may not be included. This has been experienced by the Ministry of Environment, and hence another encoding system was proposed by them.

2. The SME river coding system

The coding system used by the Spanish Ministry of Environment for water quality monitoring is very simple and based on a vectorial schema of rivers.

The codes name whole rivers or tributaries. Firstly the river network has to be digitised and encoded with a random code. The topology of the rivers and tributaries is taken into account creating an attribute table where four more data are present: the code of the river in which the river exits, the distance from this exit to the exit of the receiver river, and a flag (L/R) that indicates if it is a left or right tributary. Finally there is a code for the outlet point in the sea.



RIVER CODE	RECEIVER RIVER CODE	DISTANCE TO EXIT	SIDE	OUTLET POINT CODE
COD1	-	-	-	CODEX1
COD2	COD1	PK2	R	-
COD3	COD2	PK3	L	-
COD4	COD1	PK4	L	-

This system avoids the problems of the CDR system if any new river has to be added, (i.e., no codes have to be changed). The topology of the rivers is easy to use thanks to the relationships obtained from the attributes table.

This system has been adopted also for all the water quality monitoring networks.

1.10 The Erica catchment coding system

The EEA issued an Open Call for Tender in September 1996 to provide a River and Catchment Boundary Database. The UK's Institute of Hydrology (IH), Denmark's National Environmental Research Institute (NERI) and the University of Freiburg (UF) were successful in the tender and work began on European Rivers and Catchments (ERICA) in February 1997. The project report, by R.W. Flavin, et al., was delivered on completion of the service contract, in June 1998.

ERICA provided a low resolution (1:1 Million scale) digital database of rivers, canals, lakes, coastline and international boundaries for all of Europe. In addition ERICA provided a medium resolution (1:250K scale) pilot study databases concentrate on two areas - the Miño in Spain and Portugal, and the Meuse in France and Belgium. Section IV of the report describes in detail, and recommends, a systematic catchment coding system suitable for the EEA and its operations. The ETC/IW provided consultancy services to guarantee that the databases were developed in-line with the EEA's needs.

The ERICA project initially reviewed existing standards, some of which are detailed in this document, and claims to have used the best points of these. Like other coding systems, it uses nested codes.

In the ERICA system each river or tributary is identified with the following code format

MM BBB N1 N2 N3 N4,..., A

MM = a 2 digit Marine Code – to identify the sea

Marine Codes are based on the 1953 International Hydrographic Bureau system. For example North Atlantic Ocean = 23 and Irish Sea and St. George's Channel = 19.

BBB = a 3 digit Marine Border Code to identify the land-sea interface. Even numbers identify the points where the rivers meet the sea. Odd numbers are used to identify the land to sea boundaries between the river outlets. Thus at most, the most significant 499 rivers can be identified. This is done from north to south and from west to east.

N1, N2, = a series of 2 digit Nested Catchment Codes

Moving from river exit to source, the 49 most significant tributaries are identified and assigned consecutive even numbers (e.g. 2, 4, etc.). *(While it is not clear from the Erica documentation, the author assumes that less than 49 significant tributaries can be used, as will be necessary for small rivers. Example diagrams in the documentation seem to support this.)*

Each significant tributary has its own catchment. The remaining areas of the overall catchment now need to be identified. These are

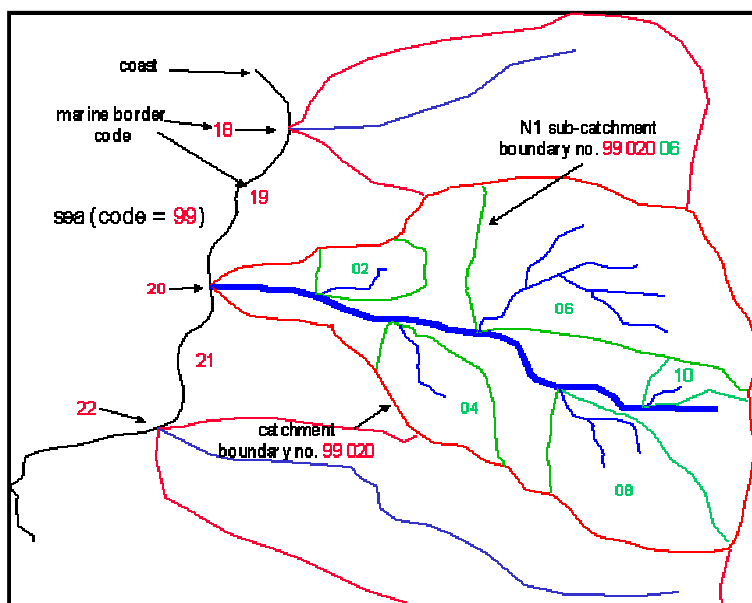


Figure A.10 Numbering the most significant tributaries

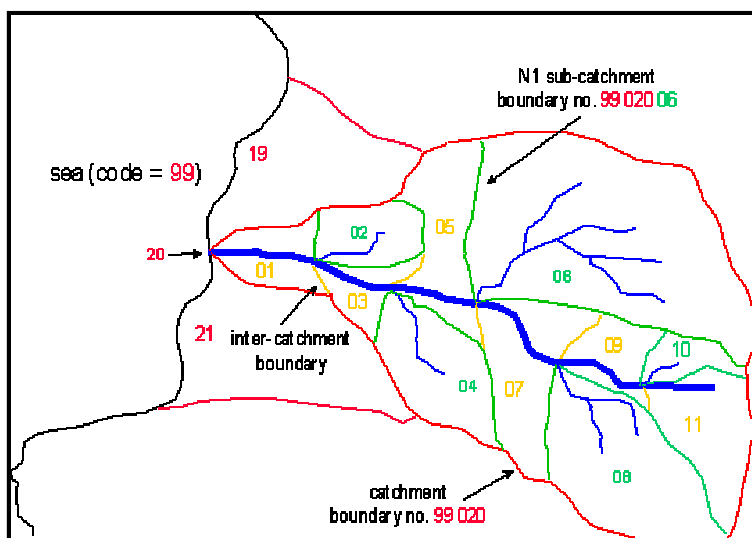


Figure A.11 Defining and numbering the inter-catchment areas

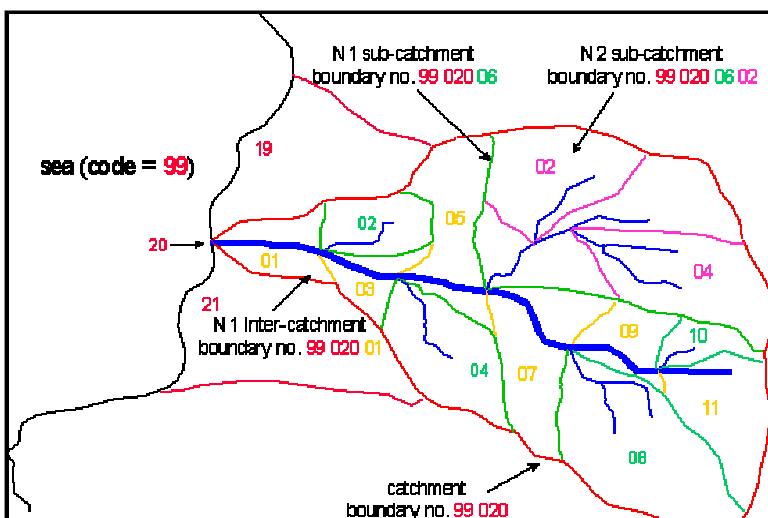


Figure A.12 Example second level tributary coding

known as inter-catchments. Each inter-catchment is now numbered using consecutive odd numbers, starting with 1 being the inter-catchment between the sea and the first significant tributary. Thus the two digits assigned to N1 enables the catchment to be subdivided into 99 sub-catchments. Each sub-catchment can then be broken down further in the same manner, by the use of N2. This nested process can continue into further levels.

Areas draining directly to sea (with diffused drainage or small rivers), will have odd numbered marine border codes and can use N1 to identify the most significant rivers, then N2 for the most significant tributaries, etc.

A = a single character Catchment Size Indicator

Finally a catchment size code is used to overcome the fact that the coding structure is not otherwise related to the size of catchments. Suggestions are made for size bands between $A = 1 \text{ km}^2$ and $N = 250,000 \text{ km}^2$.

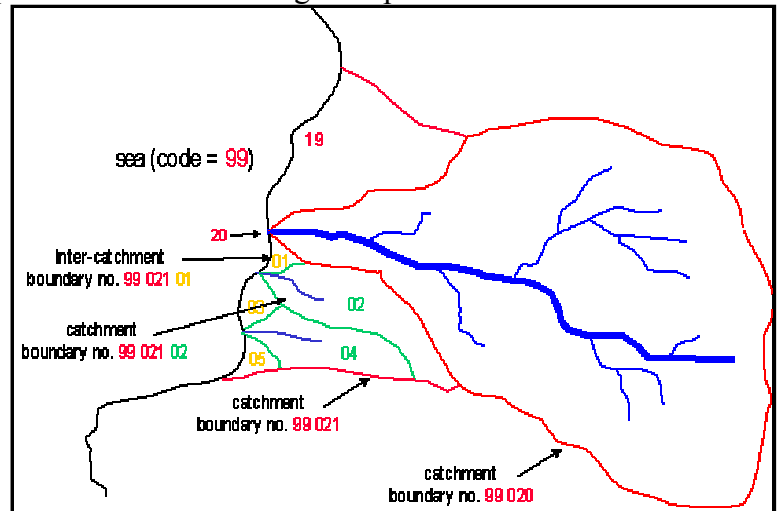


Figure A.13 Sub-division of coastal catchments

1.11 Comparison of Key River & Catchment Coding Structures

Coding System	Topology Captured	Structured Nested Coding	Easy to Understand	Extra Digit(s) assigned per level of hierarchy	Information in Digits enables direct analyses	Hydro logical catchments	No. of Nested Sub-Catchments	No. of Nested Tributary Catchments	Indicates Catchment Size	Consistent treatment of Coastal Areas	Starts with Marine Codes
Pfafstetter	Y	Y	Y	Y	Y	Y	9	4	*	Y	*
ERICA	Y	Y	Y	Y	Y	Y	99	49	Y	Y	Y
LAWA	Y	Y	Y	Y	*	Y	9	4	*	*	*
REGINE	Y	Y	*	*	*	*	33	*	*	*	*
Finish	Y	Y	*	Y	Y	Y	*	100	Y	Y	*

In addition there is the new Spanish system, which uses random coding supported by a database which maintains all topological details.

1. The Pfafstetter system has the following advantages:

- It is very easy to implement and understand.
- Areas draining in a diffused manner to marine borders are treated in a manner consistent with river coding
- It enables easy assessment of relationships between sub-catchment and river reaches based on simple numeric comparisons and detection of even or odd numbers.

And the following disadvantages:

- It has a limited number of sub-catchments (9) and tributary catchments (4) within every catchment. This gives problems with maintaining an approximate size/hierarchy relationship.

2. The ERICA coding method is effectively an extension of the Pfafstetter system. It has the following advantages:

- It caters for up to 99 sub-catchments and 49 tributary catchments, thus making it easier to maintain a relationship between catchment size and levels of coding.
- Seas and Marine borders are handled.
- Areas draining in a diffused manner to marine borders are treated in a manner consistent with river coding.
- The catchment size is indicated, independently of the nested coding.
- It enables easy assessment of relationships between sub-catchment and river reaches based on simple numeric comparisons and detection of even or odd numbers.

And following disadvantages:

- The Marine Code is based on an IHO 1953 standard, which provided a 2 digit code for marine and sea areas. This has now been replaced by an updated IHO standard, which uses an extendible decimal sub-classification code, which would be harder to implement within a catchment coding system.
- The larger number of tributaries to be identified makes it much harder to implement each level. This may be particularly so when trying to home in quickly onto a specific region.
- The two digits required, to represent 49 tributaries, means that the code is longer and harder to read at a glance.

2. The LAWA system has the following advantages:

- It is very easy to implement and understand.
- It enables easy assessment of relationships between sub-catchment and river reaches based on simple numeric comparisons and detection of even or odd numbers.

And the following disadvantages:

- Like the Pfafstetter system, it has a limited number of sub-catchments (9) and tributary catchments (4) within every catchment. This gives problems with maintaining an approximate size/hierarchy relationship.
- The treatment of the coastal areas causes inconsistencies in the coding system. Within the principal catchments, all catchments are nested within each other. In coastal areas, however, at the highest coding level, areas that are not geographically connected are lumped together (first digit 9, Figure A.5 to A.7). The second digit stands not for a sub-catchment, but for geographical neighbourhood. Thus the LAWA system is not able to provide a code of equal format in all levels and areas.

4. The REGINE coding method has the following advantages:

- the qualitative distinction between catchments with one single outlet and areas with diffuse drainage gives additional information,
- the code is economical in that it uses few digits, and
- the alphanumeric format of the code enables a high number (24) of subunits for every catchment.

And the following disadvantages:

- the two different steps of subdivision, one with the addition of a new digit and one without, make the system somewhat unclear and difficult to handle,

- the concept of hydrological areas is not as clearly defined as a hydrological catchment. Coastal areas in the system are not clearly indicated (a subunit with a 0 on the second level, e. g. 016.0 in Figure A.8, is always a coastal area in Norway, (but this would not necessarily apply to other countries).

5. The Finish coding method has the following advantages:

- it is a structured hierarchical approach,
- it has been successfully applied to numerous national databases,
- subdivision appears to be on a hydrological catchment basis
- catchment size is indicated

And the following disadvantages:

- the hierarchical structure of the code is split into two portions with different methods, a regional portion and a river network portion within that region, and the codes for each element appear to overlap,
- the codes are complex and long (18 character),
- it would be difficult to introduce additional tributaries along an already coded river, as this would upset existing coding or code structures.

6. The Spanish coding method has the following advantages:

- it is a random coding system supported by a database, therefore codes are easily assigned,
- the suggested database or an equivalent one should be a supporting component the overall GIS,
- it provides a reasonable solution where the application of structured coding is deemed to be beyond immediate resource capabilities,

And the following disadvantages:

- it relies on synchronised maintenance of random codes and database entries,
- no information is carried in the codes.

A number of issues arise out of reviewing existing coding systems.

1. How many tributaries should be assigned in each of the levels of river coding?
This issue is the most significant difference between the ERICA system and most others.
 - Using the four most significant tributaries and the associated use of one digit is easy for manual assignment and subsequent interpretation.
 - Even four tributaries may be excessive at the lowest level of catchment size. To implement it may require a detail of mapping not otherwise required. Hence it may be necessary to allow a termination, which is based on less than 4 significant tributaries and 9 sub-catchments.

- Using up to 49 significant tributaries is difficult to assign manually and subsequently interpret. A quick scan of the code will not readily reveal if only inter-catchment areas are involved as every other digit may be even (e.g. 1759235763, here '2' and '6' are even numbers but they belong to 23 and 63 which are odd). The benefits of such a system are that four tributaries may be too limited, especially when starting off at the coastal boundaries at a national or European level.
2. How much information should be built into the codes?
- The evolution of coding systems has brought about an increase in the information that can be readily determined from glancing at a code.
 - To compare catchments of similar size, it is desirable that they be coded in a similar fashion. The proposed codes generally fail to address this. The ERICA and Finish systems overcome the problem by the use of further characters to identify the catchment size.
 - It could be argued that there is no need to contain such information within codes, as it can be readily stored as attributes within databases.
 - To take this a step further, should river topology to be accurately reflected in coding structures. Such geographical connectivity can instead be stored within a GIS environment or a database environment as in the case of the Spanish system..
 - By building in some information, both manual and automated procedures are made highly efficient as demonstrated in the examples outlined in the review of the US catchment coding system.
3. Existing coding systems are somewhat weighted by the importance of coastlines to the member state involved, as can be seen in the case of German and Norwegian coding systems.

1.12 Conclusions

Regardless of which coding structure is used and how much information is contained within codes, there is an overriding need to have a common way of uniquely identifying rivers catchments, water bodies and associated features. This will be needed for international RBDs and it will be needed for general reporting.

Water bodies will be the basic management unit and hence must be ready for statistical comparisons in order to derive prioritised management plans. Common coding systems will greatly ease electronic reporting and subsequent analyses.

A balance has to be struck between 'clever' coding systems and 'simple' coding systems.

The key benefits that can be realised, from 'clever' coding systems, are

1. Easy local generation of additional Unique codes.
2. Easy checking of network connectivity and catchment relationships.
3. A consistent and systematic process.
4. Enhanced data validation capabilities.

Such codes may be assigned as part of an automated coding process where digital elevation models are used to generate the catchment boundaries down to the required levels as prescribed by the WFD.

However, it must be remembered that the primary objective is to have unique coding across Europe, such that each river reach and each sub-catchment receives its own individual identifying code in a standard format. Thus 'clever' coding is a luxury which may be abandoned in the event that it proves problematic. Where clever coding is easily achieved, then it should be deployed. This can complement the powers of GIS by providing much faster answers to simple questions. This argument is strengthened for Internet access where client side processing is provided.

Data analyses can, in any event, be left completely to GIS where for example it is possible to check on connectivity and compare catchments of similar size with similar coastal or non-coastal boundaries.