Comparison of statistical algorithms for detecting homogeneous river reaches along a longitudinal continuum

Thierry Leviandier, Adrien Alber, Florence Le Ber, Hervé Piégay

Comparison of statistical algorithms for detecting homogeneous river reaches along a longitudinal continuum VERSION ACCEPTEE 20 8 2011

Thierry Leviandier[a,*], A. Alber[b], F. Le Ber[c], H. Piégay[b]

[a]ENGEES 1 quai Koch BP 1039 F 67070 Strasbourg Cedex, France
[b]University of Lyon, UMR 5600 CNRS, Site ENS Lyon, 15 Parvis R. Descartes, 69362 Lyon cedex 07, France
[c]Laboratoire d'hydrologie et de géochimie . Université de Strasbourg, CNRS, ENGEES 1 quai Koch BP 1039 F 67070 Strasbourg Cedex, France

Abstract

Seven methods designed to delineate homogeneous river segments, belonging to four families, namely — tests of homogeneity, contrast enhancing, spatially constrained classification, and hidden Markov models — are compared, firstly on their principles, then on a case study, and on theoretical templates. These templates contain patterns found in the case study but not considered in the standard assumptions of statistical methods, such as gradients and curvilinear structures. The influence of data resolution, noise and weak satisfaction of the assumptions underlying the methods are investigated. The control of the number of reaches obtained in order to achieve meaningful comparisons is discussed. No method is found that outperforms all the others on all trials. However, the methods with sequential algorithms (keeping at order $n+1$ all breakpoints found at order n) fail more often than those running complete optimisation at any order. The Hubert-Kehagias method and Hidden Markov Models are the most successful at identifying subpatterns encapsulated within the templates. Ergodic Hidden Markov Models are, moreover, liable to exhibit transition areas.

Keywords: river segmentation; spatial organisation; GIS; river continuum

*corresponding author. Tel: 33 3 88 24 82 38; Fax: 33 3 88 24 82 84; E-mail:thierry.leviandier@engees.unistra.fr

1. Introduction

Since the pioneering work of Schumm (1977), a river channel has been considered a longitudinal continuum, its width or depth increasing downstream in response to discharge, whereas its slope and grain size decrease (Leopold and Maddock, 1953). In terms of hydraulic geometry, different authors, such as Hey (1982) and Ferguson (1986) have proposed regionally based empirical models, giving support to the 'space-time substitution' model used to highlight the channel's responses to different anthropogenic pressures, such as urbanisation or channel straightening. These pressures disrupt the longitudinal trends of the width increase in downstream channel (Gregory and Chin, 2002). Following these studies, several authors stated discontinuities within the river continuum, notably in relation to the lithological settings (Ichim and Radoane, 1990; Ferguson and Ashworth, 1991; Piégay et al., 2000a), tectonics (Schumm and Spitz, 1996; Astrade and Bravard, 1999), hillslope processes (Grant et al., 1990), riparian vegetation (Clifton, 1989), hydraulic and sedimentary effects of confluences (Rice, 1999), or human infrastructures and sea level (Gurnell, 1997). Other authors underlined the effects of these physical discontinuities on the biocenotic conditions (Minshall and Petersen, 1985; Piégay et al., 2000b; Rice et al., 2001). Perry and Schaeffer (1987) then proposed the 'discontinuum river concept' in response to the so-called 'river continuum concept' of Vannote et al. (1980).

As a consequence, new research perspectives have been opened to reconsider the factors controlling the longitudinal discontinuities and to identify them along the continuum (Benda et al., 2004; Torgersen et al., 2008). Because rivers are hierarchically organised (Frissell et al., 1986), longitudinal discontinuities can be investigated for a range of spatial scales, from the habitat to the network scale. Following Frissell et al. (1986), we refer here to the segments and reaches scales, which are stretches of river with a well-identified geomorphic structure, in between the network scale and the habitat scale. Segments are considered portions of valleys with relatively clear boundaries controlled by geological settings or bounded by tributary junctions, whereas reaches are more closely associated with homogenous geomorphic patterns, also characterised by clear boundaries (e.g., meandering, braided, straight channels). The longitudinal structure studied here is very close to the temporal univariate series, the time line being replaced by a longitudinal line. Although the longitudinal periodicity of geomorphic structures has been quite well studied — as stated by Grant et al. (1990), Hardisty (1993), Madej

(1999), and Aubry and Piégay (2001) — the detection of thresholds and homogeneous reaches has not been sufficiently developed from a scientific point of view, being mainly subjectively based, and uses visual or graphic estimates (Astrade and Bravard, 1999; Gurnell et al., 2000; Michalková et al., 2011). The moving average method has been used to smooth the local variability to highlight major thresholds on a signal with higher frequency, but it may induce cyclic effects cumulating the successive random effects (Slutzky, 1937; Bernier, 1965). A cumulated frequency curve is also a graphic mean to show the unstationarity of a given univariate series visually. Empirically, fitting a linear trend and its confidence intervals to detect stationarity is possible (Brunet-Moret, 1971). Even if such a cumulated curve is not always easy to read, one or several trends and associated thresholds can then be highlighted. Most of the available tools used to characterise longitudinal fluvial patterns are therefore univariate, whereas the questions posed in this domain are often multivariate, which is also a key issue to consider. Indeed, the detection of homogeneous segments along rivers is now becoming a challenging perspective because of the GIS layers, digital elevation model (DEM), and remote sensing data available to study such problems at a regional scale and the needs of river managers to describe river systems at the basin scale (several thousand to several hundred thousand square kilometers) for planning and targeting their actions, as demonstrated by recent publications on this topic (Beechie et al., 2006; Brenden et al., 2008). The regionalisation of synthetic geomorphic indicators should highlight how the longitudinal features control the ecological potentials and contribute to the pressure-impact models.

The aim of this contribution is therefore to inventory, describe, and compare a set of statistical methods that can be used to discretise a longitudinal continuum of a continuous variable (e.g., slope, width, depth) automatically. The hydrographic network is considered here a set of longitudinal segments/reaches separated by confluences with a focus on linear features that are disconnected hierarchically. Some of the statistical methods have already been used in such contexts, whereas others are applied to temporal series such as rainfall and discharge. The aim is therefore to assess the potentiality and the sensitivity of statistical methods according to different types of spatial distribution and resolution to detect their application domain. Section 2 presents four types of methods that are considered for comparison (seven methods with variants taken into account). The principles of algorithms are compared in section 3, and an empirical benchmark is performed in section 4. In the empirical part, the methodology, restricted to the univariate case, is presented consisting of designing synthetic templates and applying the algorithms in optimal and nonoptimal conditions on these templates and on a case study. What is learned from the tests for applications is discussed in section 5 together with the limits and potential of the univariate methodology.

2. Methods for delineation

Four types of methods are investigated, some of them with many variants, others relatively specific.

- **Homogeneity test (`HT`)** methods consist of testing a null hypothesis of the homogeneity of means. They have been mainly applied to climatic and hydrological series usually to detect one threshold (sometimes based on the Bayesian approach) (Gardner Jr, 1969; Lee and Heghinian, 1977; Buishand, 1982, 1984; Aka et al., 1996; Lubes-Niel et al., 1998; Robson et al., 1998), sometimes several thresholds (Scheffe, 1959; Hubert, 1989, 2000). Others were developed for water quality surveys (Hirsch et al., 1991). Single-threshold tests such as the Pettitt test have also been applied to detect several thresholds (Alber and Piégay, 2010). Only two of these methods, the Pettitt and Hubert methods, are described in detail here. One of the most commonly used is the Pettitt test (Pettitt, 1979; Zhang et al., 2008).

- **Contrast-enhancing (`CE`)** methods share with the `HT` methods an up-down approach to splitting, but explicitely assume heterogeneity. This type is represented only by the contrast-enhancing clustering process, proposed by Leviandier et al. (2000) and applied to the design of a rainfall stochastic process. This method detects a rupture, not by the optimisation of certain criteria, but by means of a ‘separation equation’ that yields a point satisfying certain relationships between the average intensity of the variable and the length of the subinterval on the left- and right-hand sides of the point. At least two types of separation equations were found necessary; when the equations had no root on the entire interval (no intersection of two curves), taking a point minimising the distance between the two curves was proposed to avoid stopping the procedure too early.

- **Spatially constrained clustering methods (SCC)** (Brenden et al., 2008) are ordinary clustering methods, used in the particular case of spatial objects, so that a new object in a class is preferentially aggregated to classes to which nearby objects have already been aggregated. The ultimate simplification of the procedure consists of deciding whether a point belongs to a class on its left or on its right.

- **Hidden Markov Models (HMM s)** (Rabiner and Juang, 1986) are rather different. They consider a stochastic process along the line and a random variable at each point, which produces the observed value. A reader familiar with the framework adapted to the other methods would expect that the distribution of the random variable depends on the 'class' to which the current point belongs. This is slightly more complicated as each point is not conditioned by its belonging to a class but by the pathway on which it has walked from the beginning of the line by a succession of random transitions between a small number of states. As at the end of a statistical fitting, many points have a probability 1 of being in one state and adjacent points are in the same state, it produces clusters similar to that of the other methods, with the difference that some points remain conditioned by several states with different probabilities between 0 and 1. This feature offers the possibility of exhibiting transition areas.

Two methods are considered within the HT family, one of them with two variants and two variants of the HMM method, which gives a total of seven methods.

2.1. Details on HT

As these methods are based on statistical tests, the model contains a parameter, namely the probability of a type I risk (risk of rejecting the homogeneity assumption even though it is true), which governs the detection of a stepwise shift and the number of stepwise shifts if applied several times. It must be noted that the detection of several shifts should be done with generalised tests on means (Scheffe, 1959) that is not a simple iterated application of the test for one shift. That simplification generally contains a logical contradiction: a discontinuity is sought at iteration i in a segment assumed to be homogeneous at iteration i-1.

2.1.1. Description of the Pettitt test

The sequence of random variables X_1, X_2, to X_n may have a change point. The question is to evaluate the probability that the following conditions are satisfied: the sequence of random variables X_1, X_2, to X_n has a change point at T if X_t for $t = 1$, 2, to T has a common distribution function $F_1(x)$, X_t for $t = T+1$, to n has a common distribution function $F_2(x)$ and $F1(x) \neq F_2(x)$. The null hypothesis (H_0) is defined by the stationarity of the series, i.e. no change (or $T = n$). The H_0 is tested against the alternative hypothesis H_a defined by a change. Let t be the rank and K_n the nonparametric statistic:

$$K_n = \underset{t}{Min} \sum_{i=1}^{t} \sum_{j=t+1}^{n} sign\,(X_i - X_j) \tag{1}$$

with $\text{sign}(\theta) = 1$ if $\theta > 0, 0$ if $\theta = 0, -1$ if $\theta < 0$.

Let k be the value taken by K_n on the distribution. The significance probability p associated with k is determined approximately by

$$p \sim 2exp\left(\frac{-6k^2}{n^3 + n^2}\right) \tag{2}$$

If p is inferior to the risk α defined by the operator, then H_0 is rejected and a change point is localised at the rank d for which K_N occurs. The Pettitt test has been implemented to detect several change points in a given statistical distribution by running the algorithm iteratively. Considering a given initial distribution, it is stationary if no change point is detected for a given α risk. If a change point is detected, two new subseries are generated on each side of the change point, the latter being integrated into the downstream reach. The Pettitt test runs on subseries as long as change points are detected.

2.1.2. The Hubert method and its reformulation in dynamic programming

The principle of the Hubert method (Hubert, 1989, 2000) is to optimise the approximation of a one-dimensional function by a piecewise constant function under the constraint that two adjacent segments are significantly different. The method does not violate the assumptions of tests, though considering means pairwise may be seen as an approximation of the Scheffe test.

The Hubert method was initially proposed with an algorithm of low efficiency, unable to work on large datasets. Kehagias et al. (2005) noted that the problem could be written within the dynamic programming framework and therefore be solved with a fast-running algorithm. Though the optimum at order $k+1$ cannot be derived from the optimum at order k, the idea is that the solution at order $k+1$, finishing at point t, can be written as a function of the set of solutions at order k, optimal on the length [1,s] with $s < t$. Let p depend on k such that [p,s] is the last segment in the optimal segmentation $S(k, s)$.

Let $c(s, t)$ be the cost of using the mean (or another function) instead of the observed values between points s and t and $C(k, s)$ the minimal cost at order k between 1 and s. Let v be the last point.

The Hubert method is reformulated in Eqs. (3) to (5):

$$q = \underset{s \in S}{Min} \; C(k, s) + c(s, t) \tag{3}$$

with S defined by significant changes in s and t in interval p, v

$$p(k+1, t) = q \tag{4}$$

$$C(k+1, t) = C(k, q) + c(q, t) \tag{5}$$

With appropriate initialisation, backtracking the optimal segmentation, and the rules to stop the procedure, the optimisation problem under the constraints of significant changes is quickly solved by this system.

In this paper, the Fisher-Snedecor F function used for the test was calculated with the approximation of Li and Martin (2002) using the Fortran code of the Lahey Fortran library for the χ^2 function, owned by E.J. Szondi, whose source was Bargmann and Gosh (1963).

The proper application of the test depends on the length of the segments. We also tested the simplification using a threshold on the ratio of the sum of variances divided by the whole variance (of the segment under study); as it is independent of the length, the calculation is much faster. This simplification is not fully respectful of statistical theory and may produce a different delineation. It differs from the improvement of the algorithm proposed by Gedikli et al. (2010), which involves a reduction of the number of partitions explored.

2.2. Details on contrast enhancing

The idea was to design a simple stochastic model such that the location and the height of the step of a stepwise shift are defined together — requiring a unique random number in a Montecarlo simulation not used in this paper — . Let l_1 and l_2 be the average of the variable on each side of a point of abcissa i in a segment of length n, $t = i/n$ and $r_i = l_2/l_1$.

The separation algorithm used in this study seeks a breakpoint by the following subalgorithms, in this order:

- $max(r_i)$ under the condition that $r_i > 2$, with a constraint keeping from selecting successively two neighbouring points with this subalgorithm;
- separation equations $r_i = t^2, r_i = 2(1-t)$ (the first one delivers the right location of a stepwise shift between two horizontal straight lines) and the same equations the other way round (from upstream); and
- $min(|r_i - t^2|), min(|r_i - 2(1-t)|)$ and the same the other way round

The segmentation of previously found segments is prioritised according to values of the function $n^\alpha\ l^\beta\ \sigma^\gamma$, σ being the variance within the segment. In the first applications to rainfall-runoff modeling, it was important to give some weight to the volume, thus to parameters α, β; but in the geomorphological application, these parameters are $\alpha = \beta = 0; \gamma = 1$ to put emphasis on the variance σ.

2.3. Details on `SCC`

The VAST method (valley segment affinity search technique) shares with its parent method CAST (cluster affinity search technique) its 'intent to partition a set of n objects into K groups such that objects within groups are more similar than objects in different groups' (Brenden et al. (2008) quoting Ben-Dor et al. (1999)). The algorithm uses an affinity threshold value: the larger the threshold value is, the larger (and fewer) the groups are. The VAST program (Brenden et al., 2008) includes several possibilities for the measurement of affinity and other options, but only the first one prompted by the program was used.

2.4. Hidden Markov models

The Hidden Markov models are capable of segmenting a data sequence in stationary and transient parts and to builing up a classification of the data together with the *a posteriori* probability of this classification. In an `HMM`, there is a double stochastic process (Baker, 1974):

- **the former** (a probability density function) is hidden from the observer and is defined on a sequence of states (a Markov chain);

- **the latter** is visible — it produces an observation at each slot t depending on the probability density function that is defined on the state in which the Markov chain stays at t.

Kehagias (2004) has proposed using first-order hidden Markov models (`HMM1`) for hydrological and environmental time series. In this paper, a second-order hidden Markov model (`HMM2`) was used, where the underlying state sequence is a second-order Markov chain. Therefore, the probability of a transition between two states at point t depends on the states in which the process was at point $t-1$ and $t-2$. Thus, an `HMM2` is specified by

- **a set of** N **states** called $S = \{s_1, \ldots s_N\}$;
- **a three-dimensional matrix** (a_{ijk}) over S^3:

$$\begin{aligned} a_{ijk} &= \text{Prob}(q_t = s_k / q_{t-1} = s_j, q_{t-2} = s_i) \\ &= \text{Prob}(q_t = s_k / q_{t-1} = s_j, q_{t-2} = s_i, q_{t-3} = ...) \end{aligned} \quad (6)$$

  with the constraints $\sum_{k=1}^{N} a_{ijk} = 1,\ \forall (i,j) \in [1,N]^2$ and where q_t is the current state at point t; and

- **a set of** N **discrete distributions:** $b_i(.)$ is the distribution of observations associated with the state s_i; this distribution may be parametric, nonparametric or even given by an `HMM` (Mari and Le Ber, 2006).

An `HMM1` is usually estimated by the Baum-Welch algorithm, which is related to the EM algorithm (Dempster et al., 1977). Mari et al. (1997) have shown that an `HMM2` can be estimated in the same way. The estimation is an iterative process starting with an initial model and a corpus of sequences of observations that the `HMM2` must fit. Usually, the initial model has equiprobable transition probabilities and a uniform distribution in each state. At each step, the Baum-Welch algorithm determines a new model in which the likelihood of the sequences of observation increases. Hence this estimation process converges to a local maximum, according to the maximum likelihood (`ML`) estimation criteria (Dempster et al., 1977; Mari and Schott, 2001; Mari et al., 1997). The Kullback-Leibler distance between the distributions associated with the states (Tou and Gonzales, 1974) was used to assess the final model. Two states that are too close are merged and the resulting model is retrained.

In order to describe directional processes as time series (here longitudinal series), mainly linear (or left-to-right) `HMM` models are used. In these models, once the state has been visited, it cannot be reached again. In other words, the transition probability between state i and state j, a_{ij}, is set to 0 if $i > j$. On the contrary, in ergodic models all the transitions, retrotransitions (from $i+1$ to i), as well as distant transitions (between i and $i+2$, $i+3$, etc., are authorised (Fig. 1). Such models are used, for instance, to analyse agri- Fig. 1
cultural or genomic data. Indeed, they reveal repetitive patterns, such as crop rotations or protein promoters (Le Ber et al., 2006; Eng et al., 2009). This can also be interesting for detecting similar (characterised by the same distribution) but distant river segments or reaches; therefore both linear and ergodic models are considered in the following, written respectively `lHMM` and `eHMM`.

3. Comparing the principles of the methods

3.1. Algorithms

All methods considered perform clustering, contrary to another type of classification methods that tries to recognise predefinite patterns. They differ mainly in their focus on aggregating or disaggregating. The `HMM` method is somewhat different, as it works with a fixed number of states equivalent to clusters throughout the procedure, though it is always possible to resume the procedure with another number or to merge clusters. The methods are described according to five functions usually performed by subalgorithms (Table 1). Since these methods are not necessarily dedicated only to perform delineation, some features related to their potential to deliver other information will be discussed below. These subalgorithms are:

- **Separation:** How does the method decide there is a discontinuity, or in the `HMM` method, a more or less steep change?
- **(Retro)propagation:** Once a new step has been found, is there a correction of reaches found before?
- **Merging:** Is there a procedure for merging two adjacent reaches after they have been modified if they have become too similar?
- **Stopping:** What is the criterion to stop the procedure at an acceptable (or significant from a statistical standpoint) level?

Table 1: Comparison of methods according to embedded algorithms

	Heterogeneity tests (HT) (Pettitt, Hubert)	**Contrast enhancing (CE)**	**Spatial constraint clustering (SCC)**	**Hidden Markov Model (HMM)**
Separation	Extremum of statistical criterion	Separation equation	Similarity criterion	Likelihood maximum on the entire data set
(Retro)-propagation	No	Possible gene-ralisation	Yes	
Merging	No	Possible gene-ralisation	Similarity criterion	Similarity criterion
Stopping or control of output resolution	Statistical risk or predefinite max. size	Predefinite max. size or failure of separation	Criterion based on intra- and inter-variances or threshold on similarity	Maximum number of iterations
Prioriti-sation	Same as for separation (error minimisation in Hubert method)	Extremum on subintervals of a statistical criterion	Many variants	Irrelevant

- **Prioritisation:** When the procedure has been performed up to some order, what criterion should be used to choose the next reach to be tested for splitting (irrelevant for the HMM method)?

Table 1 presents a synthesis of the methods characterised by the subalgorithms embedded. Complementary comments are given hereafter.

- **The column on HT methods** is divided into two, because of the alternative of stopping either on the number of reaches or on a statistical test.

- **The column for the HMM method** has a two-line cell and a cell labelled irrelevant, which confirms that the method is atypical.

- **(Retro)propagation and merging** were attributed to CE methods but not to HT methods. It would be contradictory to the principle, though it would be possible to do so for any method. For the CE method, it seems more acceptable; but in this case, with the loss of the separation equations, the possibility of a stochastic model of subintervals built only on the separation points is also lost (because the variable under study is no longer determined by the points of separation).

3.2. Other features

3.2.1. Parameterisation

The methods under study have generally one or several parameters, mainly in the stopping criterion, which controls the result and particularly the number of reaches found. This will be discussed in the empirical comparison, section 4.4.

3.2.2. Multidimensional extension

The longitudinal approach is often associated with geomorphic typologies. As shown by previous authors (Mosley, 1987; Schmitt et al., 2007), geomorphic typology is based on a set of variables to try to summarise the complexity, often based on exploratory multivariate statistics. Gurnell (1997) studied the hydraulic thresholds along a fluvial continuum using a discriminating factorial analysis to test whether a statistical difference is observed among four reaches initially distinguished according to a set of criteria such as specific limits (the tidal influence, infrastructures) but also geomorphic characteristic (mean depth, mean width, wetted section). It may therefore be interesting to directly segment a river using several variables rather than to combine independent delineations or to modify a univariate one in a second step. Though a benchmark of the methods on multidimension variables is not within the scope of this paper, their potential for generalisation must not be overlooked. The `HT` methods, based on statistical tests on one-dimensional distributions, are the most difficult to generalise. In the `CE` method, the equations of separation are clearly attached to dimension 1, and the choice of the next segment candidate to be divided is not difficult to generalise. The easiest generalisation is probably to use separation equations on a one-dimensional multifactorial component. In their principles, the `SCC` and `HMM` methods are able to deal with multidimensional variables.

3.2.3. Extension to a river network

This question is also related to the need for segmentation for river typology. All the methods can be applied separately on different tributaries of a river network. The point is to retain a reasonable number of reach classes (or to describe them by a reasonable number of states). Only `SCC` and ergodic `HMM` methods are able to do this. For the other methods, the classification must be done in the second step. The `SCC` and ergodic `HMM` are moreover able to take into account links and nodes.

4. Empirical comparison on real and theoretical cases

4.1. A real case study

We suggest exploring here the longitudinal pattern of the active channel width (i.e., combined unvegetated bars and low-flow channel width), which is a key parameter for delineating geomorphic reaches defined by a homogeneous planform and considered a good proxy of ecological assemblages (Frissell et al., 1986; Thorp et al., 2006; Poole, 2010). A braided reach is very different in terms of habitat template from a meandering or a straight reach (Amoros and Petts, 1993; Thorp et al., 2006), and channel width is the primary factor discriminating planform on rivers with active bedload transport and bar development. The active channel width is therefore a good parameter for testing such statistical techniques because its longitudinal pattern can be strongly segmented with homogeneous reaches associated with natural and human controlled discontinuities (e.g., change in sediment regime at confluences, valley morphology, diking, damming, etc.). These spatial structures are usually superimposed on continuums that are also characterised by a general increase in channel width downstream as a result of the discharge increase. The continuum of the Drôme River (SE France), a complex planform gravel-bed that is already well known in terms of geomorphology (see Aubry and Piégay, 2001; Alber and Piégay, 2010) was chosen to explore the behaviour of the algorithms. This case is used to run all the methods, but also to help design theoretical templates. First of all, this continuum was disaggregated within elementary 25-m-long spatial units (generally plotted with a 100-m-long step), providing information at a finer scale than the one corresponding to the geomorphic reaches, as explained by the previous authors. The data used in this paper are every tenth of these segments, to have reasonable computing time with the slowest methods.

4.2. Testing the methods

Theoretical patterns are useful to exhibit differences between methods and to assess their capacity to deal with various features of spatial stuctures. However, the choice of such patterns is not easy and must take into account the structures encountered in the real world and the underlying models. Patterns too close to theoretical models will probably be easily recognised by the corresponding methods and were not representative of real problems. On the other hand, it makes no sense to check a method against a case obviously in contradiction with its assumption. The necessarily intermediate method is an abstraction of real cases, using mainly theoretical patterns with certain testing aims. Lubes-Niel et al. (1998) conducted a comparable study on hydrometeorological data, assessing the power and robustness of statistical tests on data that violated the test assumptions. Our aim differs in that it introduces these patterns instead of alternative statistical assumptions.
Fig. 2 provides examples of longitudinal patterns of active channel widths. When applied to the Drôme main stem between km 0 and 75 (nearly 26.5 km upstream from the Rhône confluence) for which the mean width is 60.1 m ($\pm$ 52 m) ranging from 5 to 330 m (Fig. 2), different observations can Fig. 2
be made: (i) the longitudinal structure is characterised by a rough gradient with the width increasing by 1.07 m/km. This value is a mean calculated from the linear trend linking the width per segment with the longitudinal distance. Nevertheless, this series is also characterised by local peaks such as at km 30 or 54. (ii) Plateaus are also common longitudinal structures (e.g., homogeneous reaches in terms of mean channel width) separated by stepwise

shifts or gradients (such as km 7-12, 13-20 or 70-74). However, plateaus are not dominant and can have high inner variability. Transient structures are frequent and succeed one another downstream. We then expect to distinguish segments with a signal characterised by smoothed hills and valleys (understood on the graphs of any variable, not in real geographical space). The focus on km 20-23.5 (Fig. 2B) shows a clear step separating a long narrow reach with an abrupt widening (km 21.4). The focus on km 61-70 shows a local variability where it is difficult to distinguish homogeneous reaches, possibly thresholds at km 63.8 or 65 (Fig. 2C). The different methods were applied to the entire watercourse and plotted separately from km 0 to 75 (Fig. 3) and from km 75 to 101 (Fig.4) with a zoom from km 81 to 87.5 showing the transition areas found by the `eHMM` method: Fig. 3,4

- Only one threshold is detected by all the methods distinguishing clear separate reaches in terms of mean width and inner reach variability (A on Fig. 3), F on Fig. 4 missed only by the `CE` model.

- The number of reaches, at this stage not very closely controlled, varies from 6 (`lHMM`) to 12 (`SCC`) and 15 for `eHMM`, but with repetition in the last case.

- For a moderately large number of states, the `eHMM` graphs are made easier to read by considering areas with high frequency change as transition areas (Fig. 5). Fig. 5

- Some methods provide a wide range of reach lengths with very short lengths (Pettitt, `CE`, `SCC` methods), whereas others provide reaches with more homogeneous lengths (Hubert, `eHMM`, `lHMM`).

- The aggregation patterns of `lHMM` and Hubert are very close as are those of the Pettitt and `CE` methods. The `SCC` is more specific and also more detailed than the others.

- `eHMM` provides a good image in the sense that it detected all the thresholds detected by one method or another with fairly homogeneous reaches in terms of length.

- The Hubert test is also fairly efficient, separating the different main structures.

- Its simplified version differs only by a minor undetected threshold (G on Fig. 4).

- The wide reaches corresponding to braided sections are not well detected. Peak B (see Fig. 3) is not detected by the Pettitt or the `CE` method, it is fairly narrow for Hubert and `lHMM` compared to `SCC` and `eHMM`. Inversely, peak C is well detected by the Pettitt tests, suggesting a long reach similar to `lHMM`, whereas the `SCC` method does not detect any reach. `CE`, Hubert, and `lHMM` suggest different length.

- Minor thresholds (D, E, or F) are again variously detected.

4.3. Choosing typical theoretical templates

This comparison demonstrates the need to better understand the effect of local variability and not only the difference in means, which may also be linked to the resolution of the elementary objects (e.g., the number of individuals in the series), but also the position in the series. In addition, the behaviour of the methods needs to be tested with curvilinear structures to better understand the threshold variability between methods.
The methods under study are generally and originally dedicated to identify stepwise shifts. The behaviour of the methods, when reaches have more curvilinear features or transitory reaches are gradual, is less well understood. Consequently, the patterns (Fig. 6 and Table 2) are built with stepwise Fig. 6
shifts, but transition segments are introduced, using gradients or pieces of sine functions (on less than a wavelength). As the methods may be sensitive to structures at different scales, the ability to recognise a particular elementary pattern may depend on other nearby patterns. A short step with a gradient in template 1 and with a long step in template 2 is intended to test this hypothesis.
The expected output from the comparisons are:

- assessment of parameters (thresholds) giving robustness to the detection;

- control of the number of reaches; and

- (optionally) detection of reaches with similar properties (here width).

Table 2: comprehensive description of templates (with abscissas of points for resolution 1)

template's number	Description		definition
1	**With gradient**		Plateaus =400 on intervals (1,21),(87,91),(97,108); =600 on (22,41) ; =0 on (63,86),(92,96); negative slope between 42 and 62
2	**Short and long steps**		Plateaus =100 on (1,50), (61,70); =200 on (51,60), (71,200) ; =220 on (201,300)
3	**Increasing amplitude**[b]	z=l	Plateau after l=250
4[a]	**Symmetry**[b]	z=100[c]	Plateaus up to l=33 and from l=347

a) For three methods only, see Fig. 13
b) Sine function: $s = 200 + z\ sin\left(\frac{l-150}{25}\right)/sin(4)$
c) An oscillation of 1.5 wavelength between plateaus (not used in this paper) is obtained with z = 100, truncated at l=50 and l=250

4.4. Condition for a meaningful comparison on theoretical cases

Choosing meaningful and fair criteria for comparison is not the least difficulty. If the methods are compared with the criterion used to optimise one and only one of them, there is no uncertainty on the benchmark; but to avoid more subtle bias, comparing methods at similar levels of aggregation is recommended, at least in the first step. The practical application may constrain the size or the number of the segments to find. Generally, each method has a parameter governing this number.

A comprehensive study of the sensitivity of all the methods to their parameters is not required to assess the influence of the parameter governing the number of reaches obtained. The number of reaches happens to be directly the parameter of several methods (`CE`, Hubert, `HMM`), and in the case of the `HT` methods, it is easy to run the method for a level of risk that accepts a large number of reaches and to save and plot the number of reaches as a function of the parameter. The case of the `SCC` method is different, as plotting the relationship between the number of reaches and the affinity threshold, which is the relevant parameter, requires running the program for many values of the affinity threshold.

Original data have a given resolution (possibly dependent on the objective of the acquisition), but may be aggregated at a lower resolution and the methods may be sensitive to this resolution. The method may also be sensitive to high frequency variations, which are not within the scope of the delineation. In synthetic data, these variations are simulated by a noise added to the raw templates in Fig. 6. The noise used in the following was drawn from a uniform distribution, which is not the type of distribution assumed in the underlying models (but we are not checking the identification of a model against data simulated with this model). Patterns with and without local variability are presented in Fig. 7 to 12, knowing that the templates without local variability may be too far from real data and from underlying models.

4.5. Comparisons of methods applied on theoretical cases

As a consequence of the pitfalls mentioned in section 4.4, three testing issues are considered: (i) the sensitivity of algorithms to parameters controlling the number of reaches, (ii) the sensivity to resolution and inner variability, and (iii) the resilience of the methods against the introduction of patterns (such as gradients and hills) not supported by the stochastic models underlying the methods. Note that these three issues may interfere and that the same result may be looked at from different points of view.

4.5.1. Sensitivity to parameters controlling the number of reaches

The study was conducted for the three resolutions (0.5, 1, 2) for the Pettitt method (Fig. 7). The number of reaches is nearly linear with the Fig. 7
parameter $\log(\alpha)$ and highly dependent on the resolution, which implies that it is far from its original meaning in the statistical test. The gradient is represented by very unequal steps. The last narrow step is the first one to be missed when the number of reaches decreases. For the `SCC` method, as the computing time is very long, it was done only for resolution 0.5. Fig. 8 Fig. 8
shows a very strong nonlinearity, which means that it is difficult to fit a given number of reaches. Moreover, a number of major steps are missed until a large number of reaches is obtained.

4.5.2. Sensitivity to data resolution and to inner variability

The results are also dependent on noise. Fig. 9 to 11 show the results of Fig.9
six methods out of seven, as the linear `lHMM` is presented for the case without noise, on the four templates, and the `eHMM` for the case with noise, knowing that the identification algorithm of the latter sometimes fails without noise.

Both variants of the Hubert-Kehagias model are represented, which are very similar. The first template is presented in Fig. 9 for resolution 2 and in Fig. 10 for resolution 0.5. As all the methods have problems capturing the pattern with an overly low resolution (length step 2), we studied the other templates only with the highest resolution (length pace 0.5). The second template (Fig. 11) shows that the Pettitt and contrast-enhancing methods do not detect the small last step, while the other methods see it and detect a short reach before the step that was generated by the random process. The results of the third template in Fig. 12 show that noise generally improves the method's performance (except for `SCC`). The Hubert method, simplified or not, and the `HMM` and `SCC` methods without noise yield similar results; while the Pettit and contrast-enhancing methods tend to concentrate more reaches in areas of greater variation. On the real case, the sequential methods (Pettitt and contrast-enhancing) miss the first peak. The others obtain similar results. As for real cases, for templates with noise and for a large number of states, the `eHMM` model can yield a very large number of very small reaches with the same mean.

Fig. 10

Fig. 11

Fig. 12

4.5.3. Performance on gradients and curvilinear profiles

The Pettit and contrast-enhancing models appear to be error-prone in the presence of a gradient (Fig. 10). Contrary to other methods, they split the gradient into very unequal segments, so that it would be impossible to design a post-treatment to identify a gradient. The other methods, though able to 'recognise' the gradient, may be disturbed in another way: the `SCC` and `lHMM` models (forced with a small number of states) do not see the narrow step on the right-hand side of the profile. In template 3, though it is difficult to say what should be the best delineation, the results seem better with noise than without, except for the `SCC` method. In the case without noise, again a separation exists between the Pettit and contrast-enhancing methods on one hand and the other methods on the other hand, with a longer segment delineated at the top of the hill in the first group.

4.5.4. Treatment of longitudinal repeatability and symmetry

The point of breaking the symmetry in original data is not that important in the real application, as symmetry does not exist in real river networks. Even if the focus is not on the identification of periodic structures, it can be expected that the same pattern should be recognised when repeated along the river. Fig. 13 shows the results on a symmetrical pattern, which is not developed for all methods. Sequential methods such as HT are unable to provide symmetrical segmentations from symmetrical patterns (Fig. 13-left). Fig. 13 (centre) shows that the nonsequential Hubert method (and SCC, not represented) may also be affected by side effects (the symmetrical sine function is not exactly centred). Fig. 13 (right) shows the CE method, which is the only one to separate the increasing and decreasing parts of a function, which may be valuable for hydraulics issues. Fig. 13

5. Discussion

The interpretation of the comparisons to help select a method is conventionally based on theoretical considerations, empirical results on theoretical or real cases, and computing times. This latest point is not discussed here as we used codes under different systems and written in different languages. The theoretical question is the nature of the underlying model: continuous or discontinuous, periodic or aperiodic, univariate or multivariate.

At the hydrographic network scale, scientific debate continues about the existence and type of discontinuities and homogeneous reaches. Some authors consider that discontinuities are arbitrary, the river course being mainly characterised by a transitory form. This is notably the case between channel patterns, given that the river often passes progressively from a truly braided pattern to a truly meandering one with transitory hybrid patterns between the two. This is the essence of the so-called geomorphic continuum. Other authors considered that discrete structures segment the continuum and considered this continuum a set of longitudinal homogeneous segments/reaches with distinct properties. In this debate, the hydrographic networks, or at least long segments of several tens of kilometers are recognised as complex features characterised by gradual changes downstream or distinct homogeneous reaches depending on the observation scale as well as the indicators selected. Amongst the different geomorphic variables, the average channel slope is most often a continuous variable, whereas the average channel width is more frequently characterised by abrupt discontinuities.

Whether we need an exclusive separation of reaches or accept transition areas will influence the choice of the method. The tests done here confirm that the longitudinal patterns are complex, combining plateaus, gradual transitions, clearer steps, local peaks, and period structures of different amplitudes and frequencies. It is therefore difficult to be confident with a single segmentation method and comparisons between them are useful to distinguish robust discontinuities and others that are less significant.

Four key comments emerge from this comparison: (i) Only part of the meaningful geomorphic pattern is seen if we consider only one variable. We must consider this type of approach as a first step in the geomorphic characterisation, and this variable must be selected carefully because its segmentation will have consequences on the calculation of the other variables. We consider the active channel or the floodplain widths as good preliminary geomorphic variables for establishing a first clustering because they support major planimetric discontinuities and are then fairly integrative whereas the slope may show for example more gradual evolution longitudinally. (ii) The longitudinal structures are dynamic in time and are adjusted at a different time scale, but it is difficult to infer these dynamics from a snapshot of longitudinal structures. Multitemporal series of a single variable should also be consistent to distinguish permanent structures from transitory ones in time. (iii) We still know little about scaling, how the different structures are nested within each other. Following Frissel et al. (1986), the conceptual framework is clarified, but it is not validated by data. It is still unclear whether the scale-dependent structures are nested or partly independent so that chaotic organisation may prevail over nested organisation. (iv) The signal results from different drivers so that its sequencing based on a single statistical procedure (threshold, periodicity, or gradient detection) is a significant simplification; but we first need to better understand the scale organisation to reconsider which statistical tools could be appropriate, whether or not they are combined.

The identification of periodic structures is hindered by the multiplicity of spatial signatures (monotone, periodic, chaotic, variously stationary, etc. of geomorphic variables, the scalar dependence of fluvial forms and their spatial variations, and the unclear consequences of measurement errors in data.

Some periodic structures exist within some of the reaches because of hydraulic processes structuring channel forms, notably along sections characterised by sequences of pools and riffles (Montgomery and Buffington, 1997), which may provide inner disrupting structures complexifying the detections. Several statistical tools, such as spectral analysis or auto-correlation, usually used to describe temporal organisation, have already been applied to characterise longitudinal structures along river continuum, notably to highlight alternating hydraulic features. The Fourier transform has been used to show the longitudinal organisation of the water velocity framework (Hardisty, 1993) and the channel width and slope (Nakamura and Swanson, 1993). Spatial auto-correlation functions, Geary's c and Moran's I mainly, more rarely non-ergodic correlation and covariance, have been applied to illustrate the spatial regularity of hydraulic structures or fluvial forms downstream (Aubry and Piégay, 2001). Madej (1999) calculated the Moran's I to characterise the longitudinal regularity of pool-riffle sequences along a long profile and the autocorrelation lag as a scalar index (frequency of changes in geomorphic features). The autocorrelation functions were also used to evaluate the periodicity of hydraulic parameters (grain and form roughness), the index being introduced in hydraulic formulae to predict mean flow velocity (Clifford et al., 1992). The longitudinal structure has also been studied in terms of the probability of the occurrence of geomorphic features (pool, cascades, rapids, riffles, etc.) using the Markov chain (Grant et al., 1990). The aim of periodicity detection, different from ours explained in the introduction, is then to identify the frequency of a given facies sequence (the pool-riffle sequence for example) using a transitional probability matrix where each cell corresponds to the probability that a facies can follow another one downstream. The size effect has also been considered in such studies, based notably on the fractal analysis at a hydrographic network scale (Gao and Xia, 1996).

Periodic structures were not considered the subject of this study but only as disturbances likely to be encountered when focusing on segment/reach scale for which discrete structures with fairly clear boundaries and nonperiodic structure seems to be easily detected. Consequently, we discarded the methods dedicated to recognising periodic structures because we assume they were prevailing at a local scale but are not encountered at the scale of the river reaches. Nevertheless, this comparison of algorithms showed that sinusoidal structures are frequent, sometimes periodic, and that establishing clear differences in scale levels is difficult.

Beyond the undeniable result that different methods yield different results, the comparison on theoretical templates cannot settle the scientific debate but can contribute by bringing out the biases stemming from the methods. No method is absolutely superior to all the others. However, the methods that are run sequentially (the partition at order $n+1$ takes all points of the partition at order n), the Pettitt and contrast-enhancing methods, which are consequently faster, have some propension to determine segments that are too unequal in length. This suggests a flaw in the choice of the next reach to divide rather than on the detection of heterogeneity. The Hubert method, its simplified variant, and the `HMM` methods produce — with adequate parameterisation and a reasonable number of reaches — very few failures in

detecting subpatterns introduced in the templates.
The interpretation problems with the interference of parameterisation is not a methodological issue limited to a benchmark. In real application to large data sets, it is unlikely that a method might be optimal at any scale and on the whole network. So the nature of solutions found with nonoptimal parameterisation is part of the problem. In a large network, we should also have to take into account the upstream-downstream trends that appear in the observed data.
Algorithms are moderately sensitive to noise. We did not multiply the number of trials by tuning the amplitude of noise, but only studied a few cases with and without noise. The latest case is an extreme and unrealistic simplification, in contradiction with the assumptions underlying models. However, it shows that the methods generally resist this simplification with the adapted parameters and changes in patterns found. In particular, undisturbed segments with constant gradients are easily recognised as a sequence of equal subsegments, allowing post-processing; but this procontrainperty is lost when there is noise.
Let us finally come back to the 'other features' of section 3.2, parameterisation, multidimensional, and river network extension. The benchmark was done on univariate series and highlighted only the parameterisation issue. However, as stated above, delineation is often the means for drawing up a typology that can be used as a tool for sustainable river management or for further scientific investigation. In this context, the main aim is to select the appropriate variable, the one that is sufficiently integrative to provide a preliminary segmentation and to choose the appropriate algorithm to identify the proper boundaries. Transient gradients and merging of neighbouring segments can be performed at a later stage, for example, by linear correlation and clustering. As for the last one, it is also better from a geomorphic point of view because it is based on a wider set of variables.The delineation methodology must be adapted to the ultimate aim.

6. Conclusion

Four templates were designed as a trade-off between realism, simplification, and low satisfaction of the assumptions underlying the methods under study. On these templates and on real data, seven methods belonging to four families were run.
Thresholds of statistical tests and other parameters control the number of

reaches obtained. Constraining the number of reaches, whether or not it is directly a parameter of the method, is useful to obtain comparable results for different methods. However, this number may be difficult to control in the SCC method because of the possibility of a critical value of the affinity threshold parameter.

Benchmarking should be done first, allowing 'optimal' parameterisation of each individual method; secondly looking at downgraded application conditions as it is unlikely that optimisation can be done simultaneously on different parts of a hydrographic network. These downgraded parameterisations result in omitting certain patterns that are obvious from a naive point of view. This behaviour may depend on the location of the pattern in the system.

In particular, the thresholds and parameters ensuring a given number of groups depend on the resolution of the data; but even if the number of groups is forced to equal values, the partition will not be the same at different resolutions. As the meaning of statistical tests — their interpretation as the probability of exceeding a threshold — is somewhat dubious when they are applied in a context different from their fundamental hypothesis; the thresholds become mere parameters, possibly used outside their nominal values, tuned to satisfy other criteria.

It should be noted that broader assumptions are allowed by the ergodic HMM method. If delineation of homogeneous and exclusive segments is not a compulsory objective, it may be interesting to analyse the watercourse with this method.As it is able to point out patterns other than steps, such as repetition of patterns, transition areas or high frequency alternance of substructures, without increasing the degree of freedom, it allows to revisit the objectives of delineating homogeneous segments and at least to highlight its arbitrariness. In other words, the method reveals continuum and discontinuum.

Acknowledgements

The authors are grateful to Travis Brenden for help on the VAST software and to Jean-François Mari for help on the CarottAge software. The Ph.D. of Adrien Alber was supported by the French Ministry of Agriculture. The research received funding from the Water Agency and of the ZABR for Rhône network characterisation by orthorectfied photographies.

References

Aka A, Kouame B, Paturel JE, Servat E, Niel H, Masson J. Analyse statistique de l'évolution des écoulements en Côte d'Ivoire. IAHS publication 1996;:167–78.

Alber A, Piégay H. Spatial disaggregation and aggregation procedures for characterizing fluvial structures at the network-scale: application to the Rhône basin (France). Geomorphology 2010;125(4):343–60.

Amoros C, Petts GE. Hydrosystèmes fluviaux. Collection d'écologie 24. volume 24. Masson, Paris, 1993.

Astrade L, Bravard JP. Energy gradient and geomorphological processes along a river influenced by neotectonics (the Saône river, France). Geodinamica Acta 1999;12(1):1–10.

Aubry P, Piégay H. Pratique de l'analyse de l'autocorrélation spatiale en géomorphologie fluviale : définitions opératoires et tests. Géographie Physique et Quaternaire 2001;55(2):115–33.

Baker JK. Stochastic modeling for automatic speech understanding. In: Reddy D, editor. Speech Recognition. Academic Press,New-York; 1974. p. 521–42.

Bargmann R, Gosh SP. Statistical Distribution Programs for a Computer Language. IBM Research Report RC-1094. Technical Report; IBM; 1963.

Beechie T, Liermann M, Pollock M, Baker S, Davies J. Channel pattern and river-floodplain dynamics in forested mountain river systems. Geomorphology 2006;78(1-2):124–41.

Ben-Dor A, Shamir R, Yakhini Z. Clustering gene expression patterns. Journal of Computational Biology 1999;6(3-4):281–97.

Benda L, Poff N, Miller D, Dunne T, Reeves G, Pess G, Pollock M. The network dynamics hypothesis: how channel networks structure riverine habitats. BioScience 2004;54(5):413–27.

Bernier J. Sur les probabilités d'occurrence des sècheresses et des étiages. Bull Cen Rech Essais de Chatou 1965;11:3–12.

Brenden T, Wang L, Seelbach P, Clark R, Wiley M, Sparks-Jackson B. A spatially constrained clustering program for river valley segment delineation from GIS digital river networks. Environmental Modelling and Software 2008;23:638–49.

Brunet-Moret Y. Etude de l'homogénéité de séries chronologiques de précipitations annuelles par la méthode des doubles masses. Cah ORSTOM, ser Hydrol 1971;8(4):3–31.

Buishand T. Some methods for testing the homogeneity of rainfall records. Journal of Hydrology 1982;58(1-2):11–27.

Buishand T. Tests for detecting shift in the mean of hydrological time series. Journal of Hydrology 1984;73:51–69.

Clifford N, Robert A, Richards K. Estimation of flow resistance in gravel-bedded rivers: A physical explanation of the multiplier of roughness length. Earth Surface Processes and Landforms 1992;17:111–26.

Clifton C. Effects of vegetation and land use on channel morphology. In: Gresswell RE, editor. Practical approaches to riparian resource management, an educational workshop. Billings, Montana, USA; 1989. p. 121–9.

Dempster A, Laird N, Rubin D. Maximum-likelihood from incomplete data via the EM algorithm. Journal of Royal Statistic Society, Ser B (methodological) 1977;39:1–38.

Eng C, Asthana C, Aigle B, Hergalant S, Mari JF, Leblond P. A new data mining approach for the detection of bacterial promoters combining stochastic and combinatorial methods. Journal of Computational Biology 2009;16(9):1211–25.

Ferguson RI. Hydraulics and hydraulic geometry. Progress in Physical Geography 1986;10:1–31.

Ferguson RI, Ashworth PJ. Slope-induced changes in channel character along a gravel-bed stream: the allt dubhaig, scotland. Earth surface processes and landforms 1991;16:65–82.

Frissell C, Liss W, Warren C, Hurley M. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental management 1986;10(2):199–214.

Gao J, Xia Z. Fractals in physical geography. Progress in Physical Geography 1996;20:178–91.

Gardner Jr L. On detecting changes in the mean of normal variates. The Annals of Mathematical Statistics 1969;40(1):116–26.

Gedikli A, Aksoy H, Erdem Unal N, Kehagias A. Modified dynamic programming approach for offline segmentation of long hydrometeorological time series. Stochastic Environmental Research and Risk Assessment 2010;24(5):547–57.

Grant G, Swanson F, Wolman M. Pattern and origin of stepped-bed morphology in high-gradient streams, western Cascades, Oregon. Geological Society of America Bulletin 1990;102:340–52.

Gregory KJ, Chin A. Urban stream channel hazards. Area 2002;34(3):312–21.

Gurnell A. Adjustments in river channel geometry associated with hydraulic discontinuities across the fluvial-tidal transition of a regulated river. Earth Surface Processes and Landforms 1997;22:967–85.

Gurnell A, Petts G, Harris N, Ward J, Tockner K, Edwards P, Kollmann J. Large wood retention in river channels: the case of the Fiume Tagliamento, Italy. Earth Surface Processes and Landforms 2000;25(3):255–75.

Hardisty J. Time series analysis using spectral techniques: oscillatory currents. Earth Surface Processes and Landforms 1993;18:855–62.

Hey R. Gravel-bed rivers: form and processes. In: Hey R, Bathurst J, Thorne C, editors. Gravel-bed Rivers. John Wiley & Sons, Chichester, UK; 1982. p. 5–13.

Hirsch R, Alexander R, Smith R. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research 1991;27(5):803–13.

Hubert P. Segmentation des séries hydrométéorologiques : application à des séries de précipitations et de débits de l'afrique de l'ouest. Journal of Hydrology 1989;110:349–67.

Hubert P. The segmentation procedure as a tool for discrete modeling of hydrometeorological regimes. Stochastic Environmental Research and Risk Assessment 2000;14(4):297–304.

Ichim I, Radoane M. Channel sediment variability along a river: a case study of the Siret River (Romania). Earth Surface Processes and Landforms 1990;15:211–25.

Kehagias A. A hidden Markov model segmentation procedure for hydrological and environmental time series. Stochastic Environmental Research and Risk Assessment 2004;18:117–30.

Kehagias A, Nidelkou E, Petridis V. A dynamic programming segmentation procedure for hydrological and environmental time series. Stochastic Environmental Research and Risk Assessment 2005;20(1):77–94.

Le Ber F, Benoît M, Schott C, Mari JF, Mignolet C. Studying crop sequences with CarrotAge, a HMM-based data mining software. Ecological Modelling 2006;191(1):170–85.

Lee A, Heghinian S. A shift of the mean level in a sequence of independant normal random variables: a Bayesian approach. Technometrics 1977;19(4):503–6.

Leopold L, Maddock T. The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. US Geological Survey Professional Paper 1953;252. 56pp.

Leviandier T, Lavabre J, Arnaud P. Rainfall contrast enhancing clustering processes and flood analysis. Journal of Hydrology 2000;240(1-2):62–79.

Li T, Martin EB. An approximation to the F distribution using the chi-square distribution. Computational Statistics & Data Analysis 2002;40:21–6.

Lubes-Niel H, Masson J, Paturel J, Servat E. Variabilité climatique et statistiques. Etude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques. Rev Sci Eau 1998;11(3):383–408.

Madej M. Temporal and spatial variability in thalweg profiles of a gravel-bed river. Earth Surface Processes and Landforms 1999;24:1153–69.

Mari JF, Haton JP, Kriouile JP. Automatic word recognition based on second-order Hidden Markov Models. IEEE Transactions on Speech and Audio Processing 1997;5:22–5.

Mari JF, Le Ber F. Temporal and spatial data mining with second-order hidden Markov models. Soft Computing – A Fusion of Foundations, Methodologies and Applications 2006;10(5):406–14.

Mari JF, Schott R. Probabilistic and Statistical Methods in Computer Science. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

Michalková M, Piégay H, Kondolf GM, Greco S. Lateral erosion of the Sacramento River, California (1942-1999), and responses of channel and floodplain lake to human influences. Earth Surface Processes and Landforms 2011;36:257–72.

Minshall G, Petersen R. Towards a theory of macroinvertebrate community structure in stream ecosystems. Archiv für Hydrobiologie 1985;104(1):49–76.

Montgomery D, Buffington J. Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin 1997;109(5):596–611.

Mosley M. The classification and characterization of rivers. In: Richards K, editor. River Channels: Environment and Process. Blackwell, Oxford; 1987. p. 295–320.

Nakamura F, Swanson F. Effects of coarse woody debris on morphology and sediment storage of a mountain stream system in western Oregon. Earth Surface Processes and Landforms 1993;18:43–61.

Perry J, Schaeffer D. The longitudinal distribution of riverine benthos: a river dis-continuum? Hydrobiologia 1987;148(3):257–68.

Pettitt AN. A non-parametric approach of the change-point problem. Applied Statistics 1979;28(2):126–35.

Piégay H, Salvador P, Astrade L. Réflexions relatives à la variabilité spatiale de la mosaïque fluviale à l'échelle d'un tronçon. Zeitschrift für Geomorphologie 2000a;44:317–42.

Piégay H, Thévenet A, Kondolf GM, Landon N. Physical and human factors influencing fish habitat distribution along a mountain river continuum, Drôme River, France. Geographiska Annaler 2000b;82:121–36.

Poole G. Stream hydrogeomorphology as a physical science basis for advances in stream ecology. Journal of the North American Benthological Society 2010;29(1):12–25.

Rabiner L, Juang B. An introduction to hidden Markov models. IEEE ASSP Magazine 1986;3(1):4–16.

Rice S. The nature and controls on downstream fining within sedimentary links. Journal of Sedimentary Research 1999;69(1):32–9.

Rice S, Greenwood M, Joyce C. Macroinvertebrate community changes at coarse sediment recruitment points along two gravel bed rivers. Water Resources Research 2001;37(1):2793–803.

Robson A, Jones T, Reed D, Bayliss A. A study of national trend and variation in UK floods. International Journal of Climatology 1998;18(2):165–82.

Scheffe H. Tukey's test for pairwise comparisons. In: The Analysis of Variance. Wiley and Sons, Inc., New York; 1959. 477pp.

Schmitt L, Maire G, Nobelis P, Humbert J. Quantitative morphodynamic typology of rivers. A methodological study based on the French Upper Rhine basin. Earth Surface Processes and Landforms 2007;32(11):1726–46.

Schumm S. The Fluvial System. Wiley, New York, 1977. 338 pp.

Schumm S, Spitz W. Geological influences on the lower Mississippi river and its alluvial valley. Engineering Geology 1996;45(1-4):245–61.

Slutzky E. The summation of random causes as the source of cyclic processes. Econometrica: Journal of the Econometric Society 1937;5(2):105–46.

Thorp J, Thoms M, Delong M. The riverine ecosystem synthesis: biocomplexity in river networks across space and time. River Research and Applications 2006;22(2):123–47.

Torgersen C, Gresswell R, Bateman D, Burnett K. Spatial Identification of Tributary Impacts in River Networks; John Wiley & Sons Inc, New York, USA. p. 159–81.

Tou JT, Gonzales R. Pattern Recognition Principles. Addison-Wesley, Reading, MA, USA, 1974.

Vannote R, Minshall G, Cummins K, Sedell J, Cushing C. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 1980;37:130–7.

Zhang S, Lu XX, Higgitt DL, Chen C, Han J, Sun H. Recent changes of water discharge and sediment load in the Zhujiang (Pearl River) basin, China. Global and Planetary Change 2008;60(3-4):365–80.

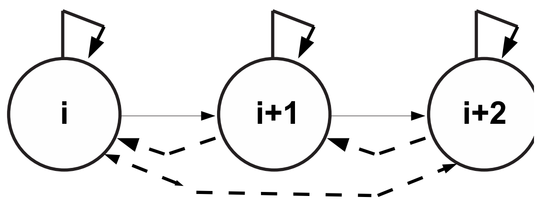

Figure 1: Topology of an ergodic HMM: the dashed lines correspond to the authorised transitions in the ergodic model vs. the linear model

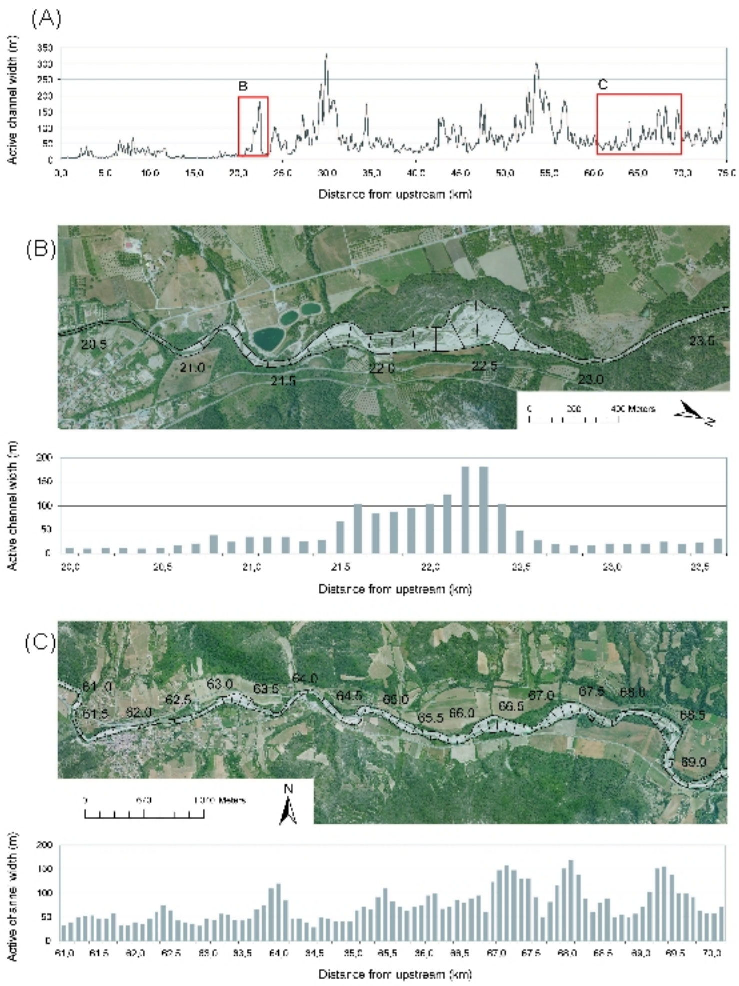

Figure 2: The Drôme River showing (A) longitudinal evolution of the active channel width showing a complex pattern with plateaus as well as hills, valleys, and gradients. Evident stepwise shifts are observed (B) as well as reaches with a high local variability and less evident shifts (C). Length of segments is 100 m

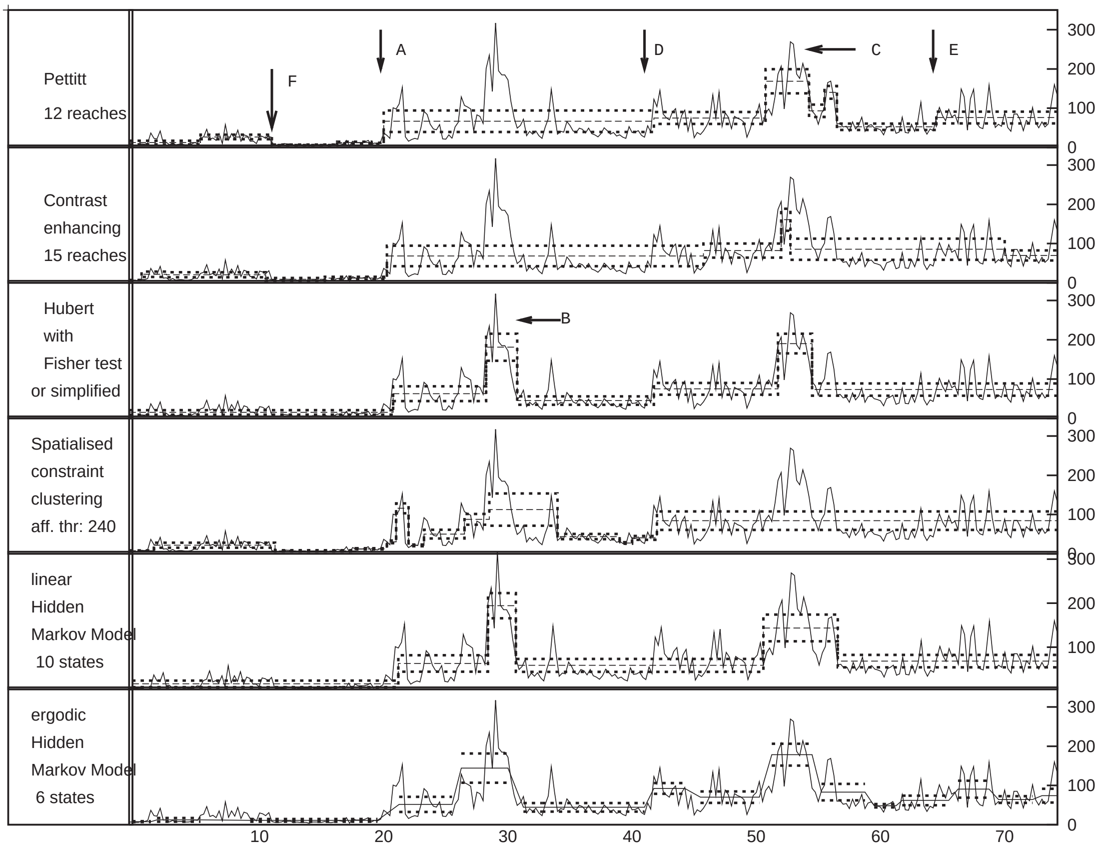

Figure 3: Drôme River, active channel width, mean, and half standard error for reaches found by all methods fitted on the whole river but represented on the upstream part corresponding to Fig. 2. For ergodic `HMM`, homogeneous sectors for no less than six points. For SCC, aff. thr. is the affinity threshold, which has the dimension of the variable.
Note that the Pettitt, `CE` and `SCC` methods are parameterised with more reaches than the Hubert and `HMM` methods

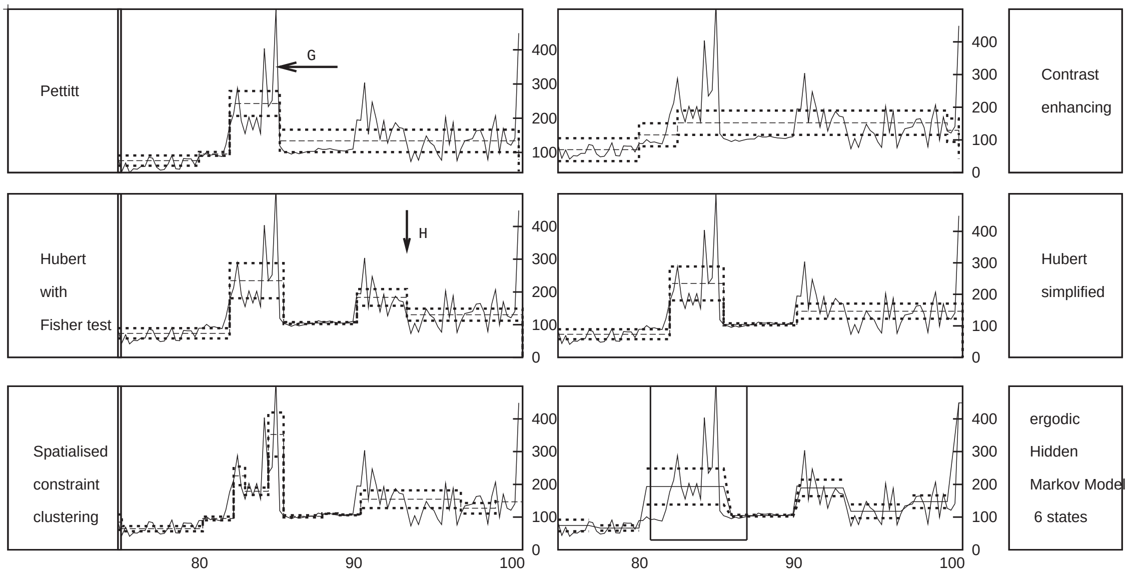

Figure 4: Part of Drôme River downstream reaches on Fig. 3. All methods with the same optimisation as on Fig. 3. See zoom of the rectangle on the `eHMM` subfigure in Fig. 5.

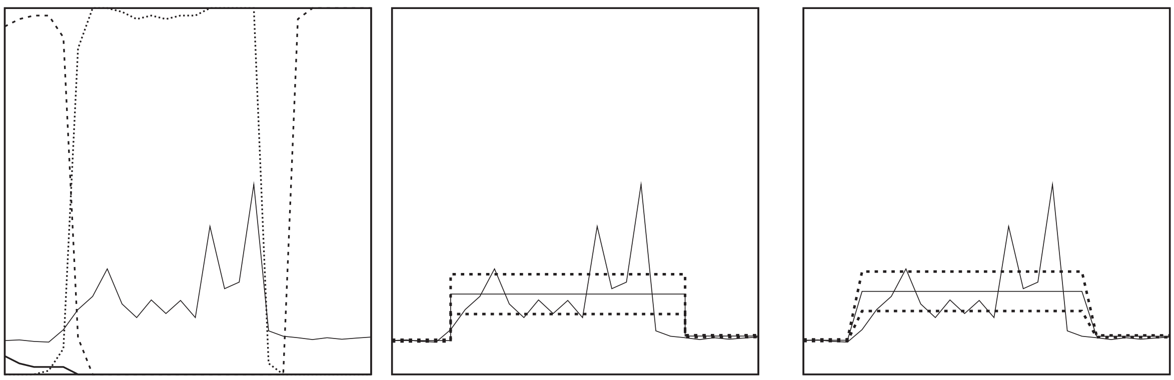

Figure 5: Zoom (km 81 to 87.5) on the Drôme River, active channel width, ergodic `HMM`, probabilities of six states. Left: probabilities of different states; centre: reaches separated on most probable state; right: transition areas for less than six adjacent points having the same most probable state.

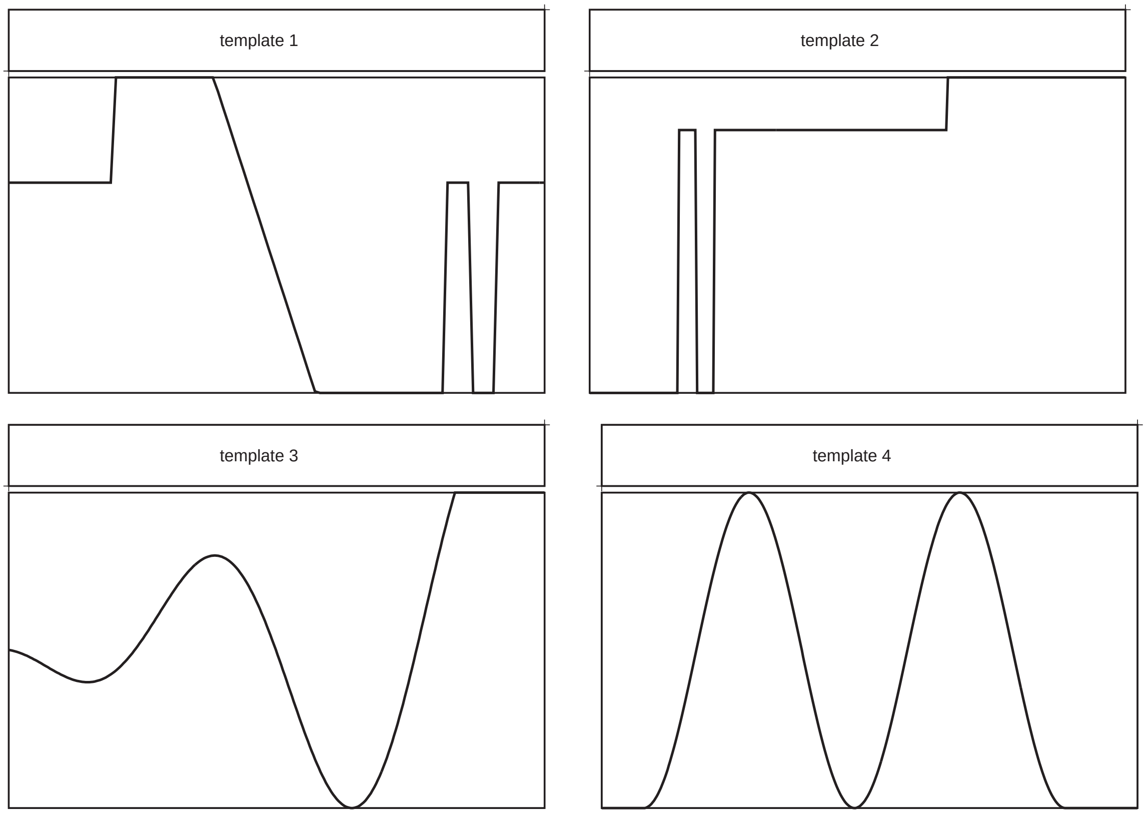

Figure 6: The four templates (without noise).

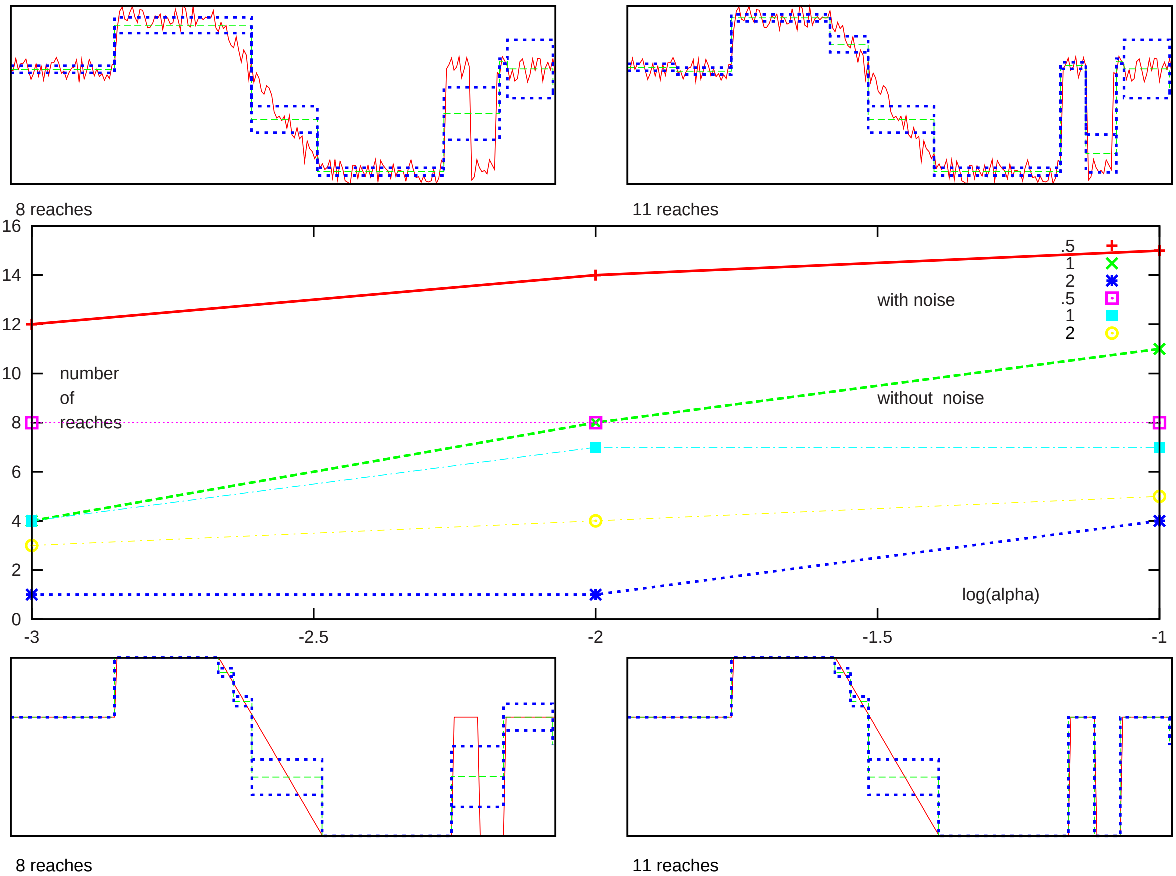

Figure 7: Sensitivity to α (probability of 2nd type risk) and resolution in the the Pettitt method, top: with noise; bottom: without noise; left: low probability of wrong acceptance of homogeneity; right: high probability. Top and bottom graphs are all for resolution 0.5.

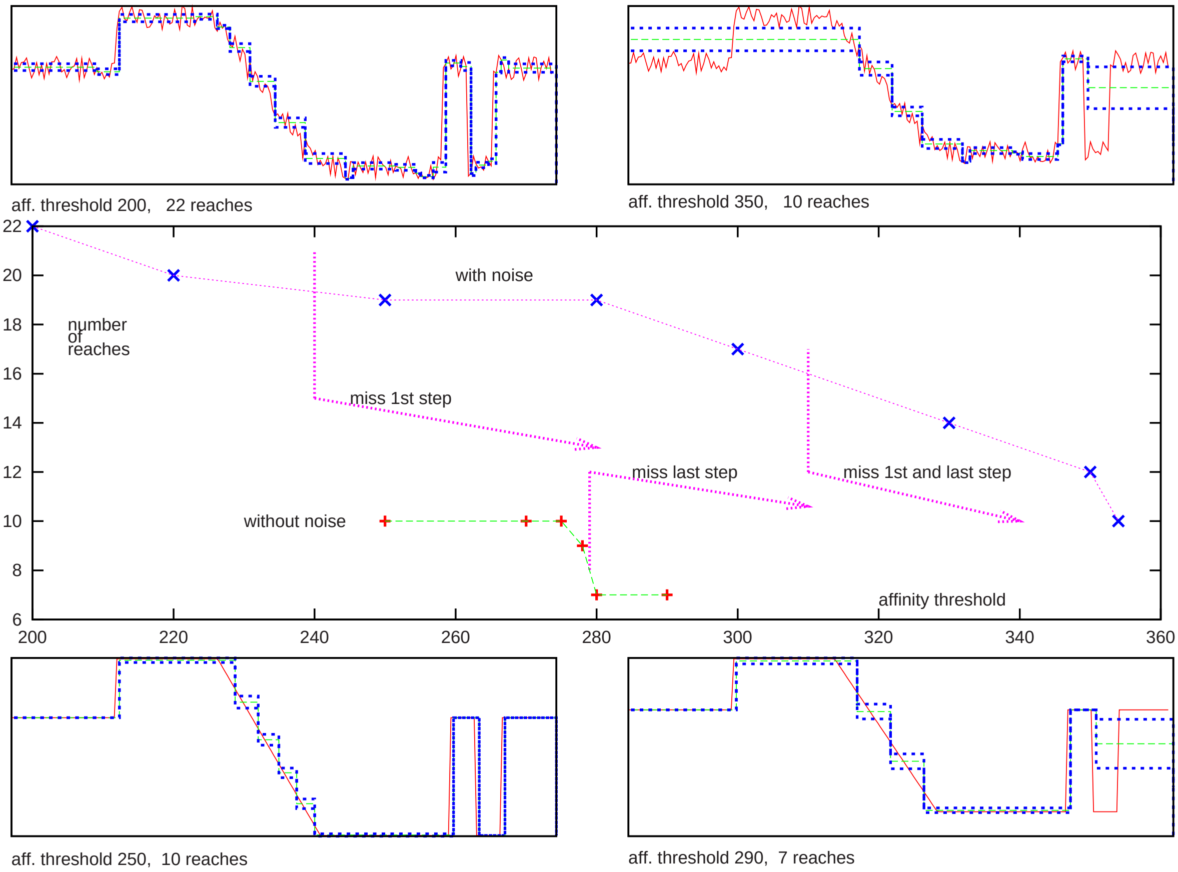

Figure 8: Sensitivity to affinity threshold in the SCC method. Top: with noise; bottom: without noise; left: low affinity threshold; right: high affinity threshold.

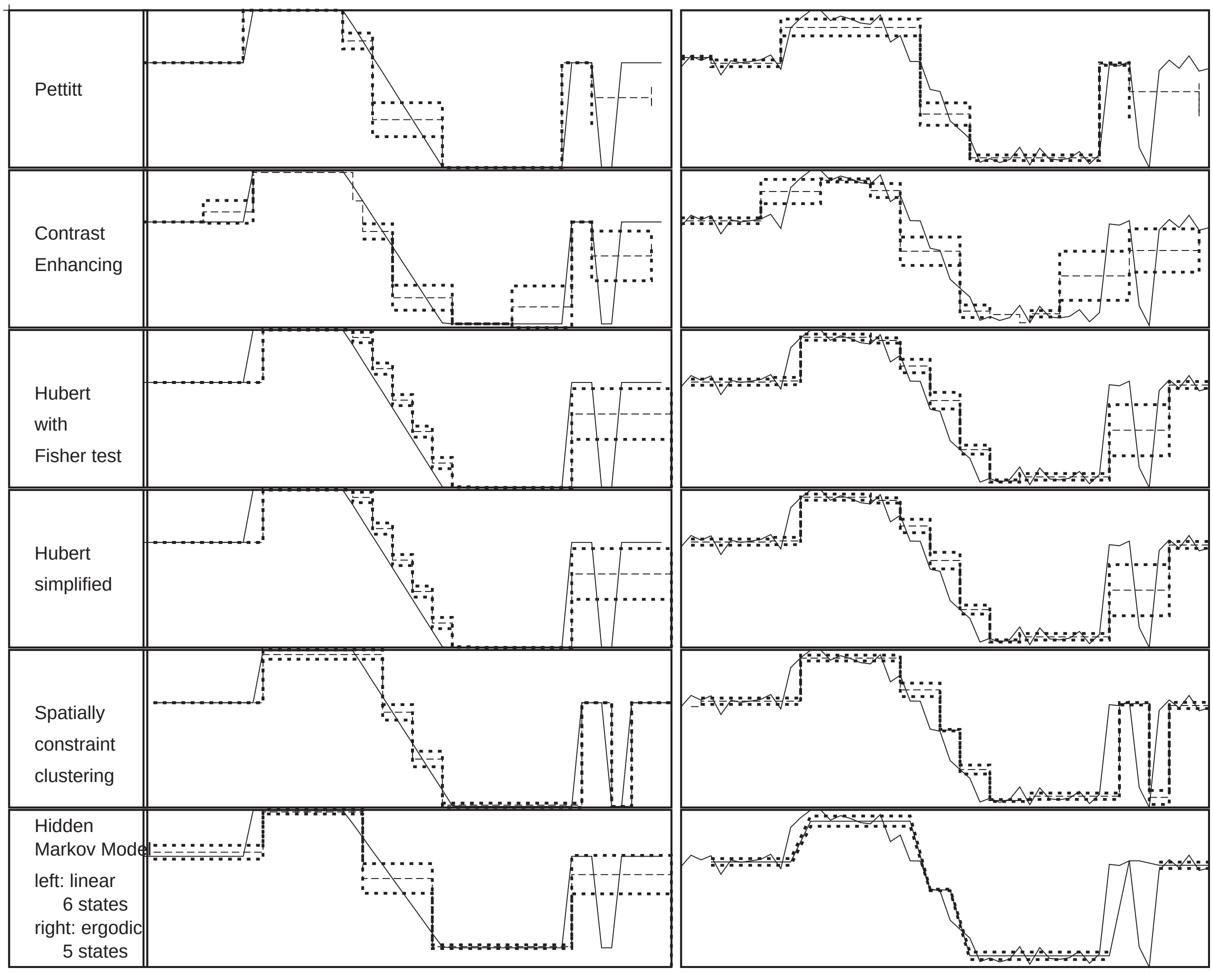

Figure 9: 1st template, resolution 2. Left: without noise; right: with noise.

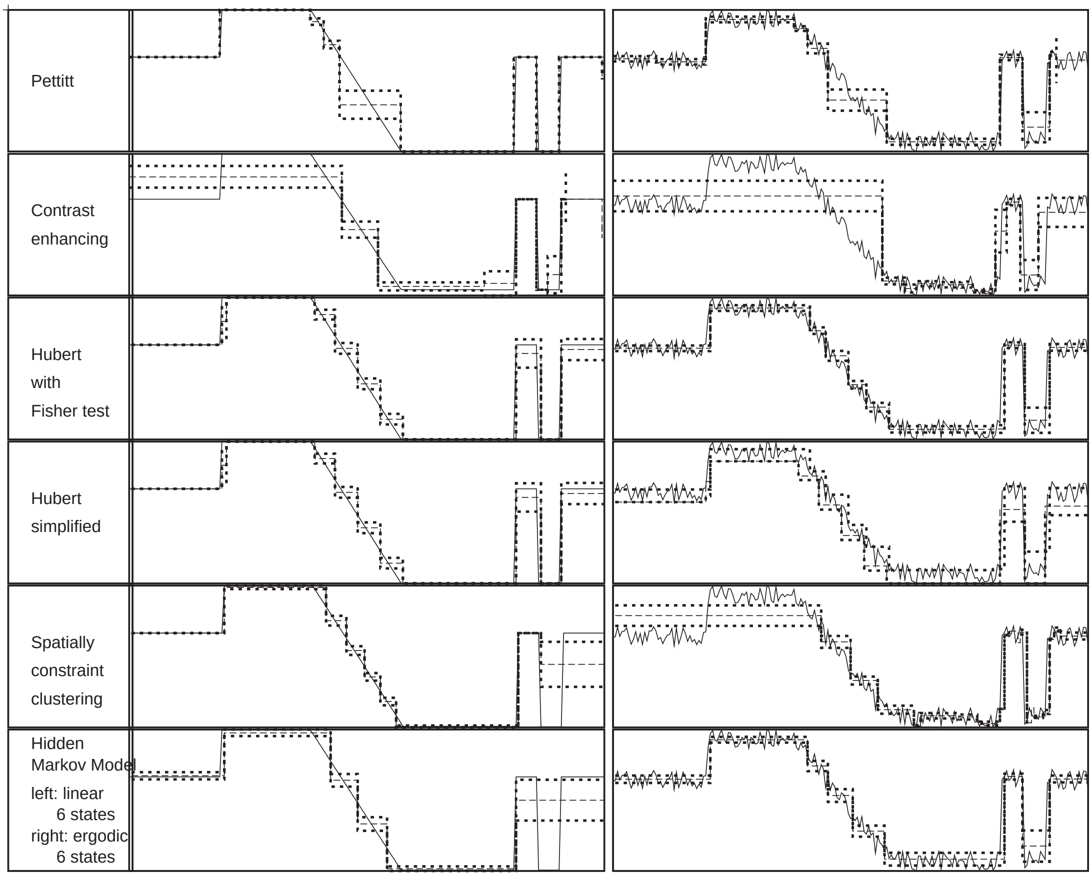

Figure 10: 1st template, resolution 0.5. Left without noise; right: with noise. The linear HMM is not optimal but presented to show the same number of states as the ergodic HMM.

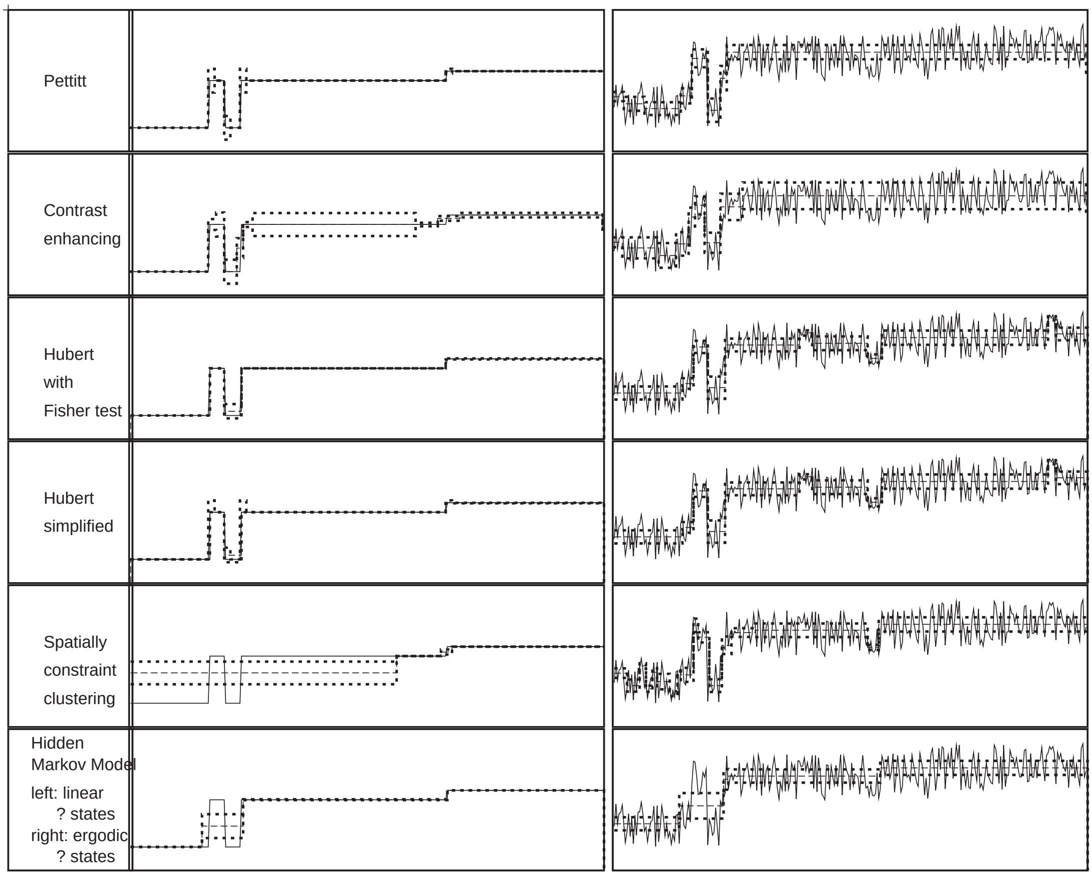

Figure 11: 2^{nd} template, with noise, all methods.

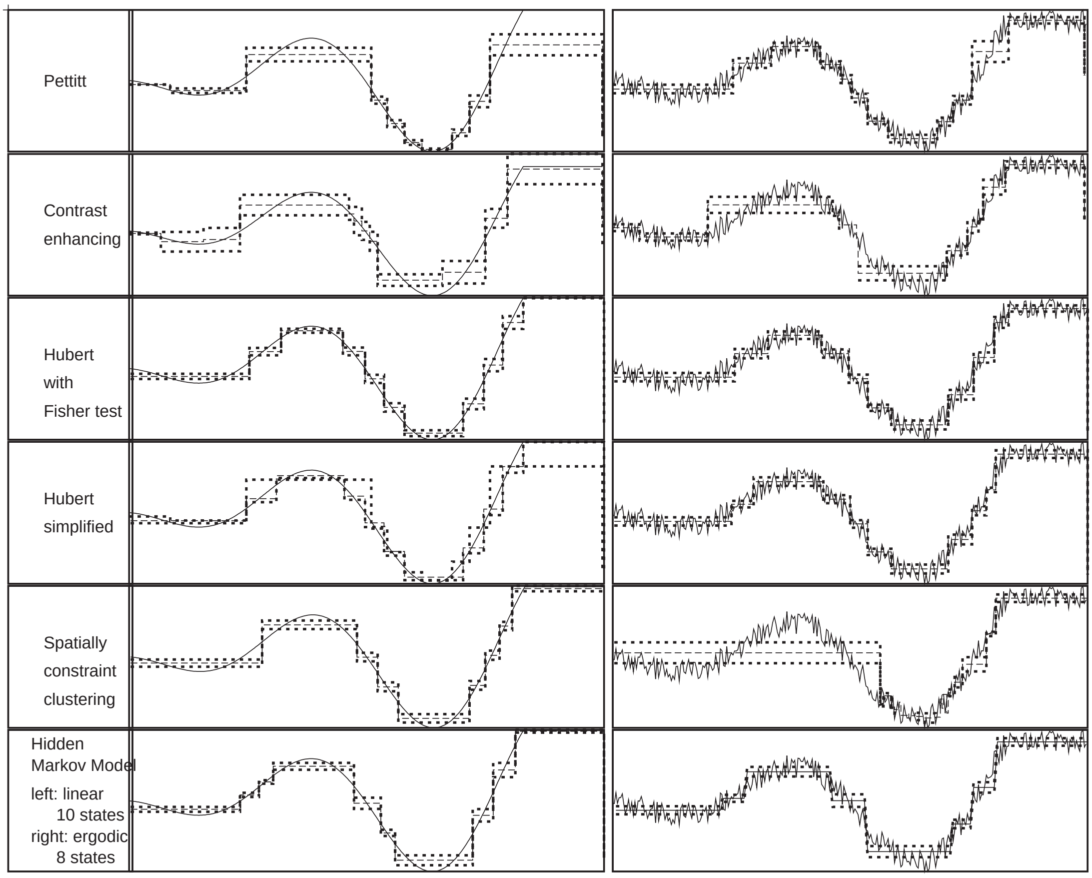

Figure 12: 3^{rd} template. Left without noise; right: with noise.

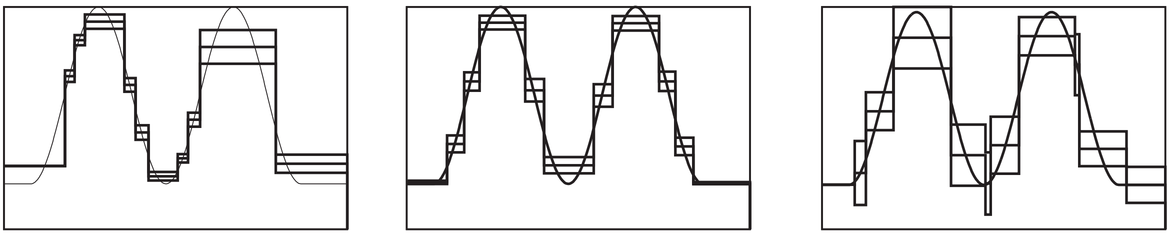

Figure 13: 4^{th} template, treatment of symmetry in different methods. Left: Pettitt; centre: Hubert(simplified); right: contrast-enhancing.