CHANGE TREES AND MUTAGRAMS FOR THE VISUALIZATION OF LOCAL CHANGES IN SEQUENCE DATA

-----------------------

Key Words: Biological sequence data, DNA sequence, jump discontinuity, nonparametric segmentation, scaling, slope change
Abstract

The analysis of local changes in sequence data is of interest for various applications such as the segmentation of DNA and other genetic sequences, sequences of climate and environmental data, financial data and event data. We propose simple graphical tools to visualize patterns of local change. The concept of mode trees, developed for the visualization of local patterns in densities (Minnotte and Scott, 1993), is adapted to visualize patterns of local change in dependency on a threshold parameter by means of a Change Tree. The additional effect of the choice of window widths motivates another graphical device, the Mutagram. We illustrate these concepts with several sets of sequence data.
1. INTRODUCTION

The statistical analysis of abrupt changes in the mean or slope of a regression function has been widely studied. The location where such a sudden change occurs is referred to as a change-point. Change-points have been studied in linear models (Kim, 1994; Bhattacharya, 1994; Jandhyala and MacNeill, 1997), non-linear models (Jandhyala and Al-Saleh, 1997; Rukhin and Vajda, 1997), generalized linear models (Braun, Braun and Müller, 2000) and non-parametric models (Hall and Titterington, 1992; Müller, 1992; Hall, Gijbels, and Kneip (1999); Darkhovski, 1994; Antoniadis, Gijbels and MacGibbon, 2000). For an overview on some developments in theory and application, we refer to the monograph on Change-point Problems edited by Carlstein, Müller and Siegmund (1994).

In this paper we propose several tools for the detection and display of change-points, based on local quasi-likelihood comparisons. The proposed methods are model free and graphical in nature. A key graph is the Change Tree. Change Trees are constructed by obtaining local quasi-likelihood fits with and without change-points at a set of potential locations. The locations are then ranked by evaluating the magnitude of the differences of quasi-deviances between fits with and without change-point at each location. The resulting ranked change-point locations are then represented in a tree graph, the root sizes indicating the size of the observed difference in quasi-likelihood between the two fits. The longest roots of the tree correspond to locations with the largest change in quasi-likelihood, the next longest roots indicate points with the next largest changes in quasi-likelihood, and so on.

A second key graph is the Mutagram which supplements Change Trees. It allows us to study variable local window sizes, referring to the window within which the local quasi-likelihoods are computed. The variation in window widths or scaling has an influence on the quasi-likelihood differences. By graphing quasi-likelihood deviance differences simultaneously for different bandwidths, the mutagram reveals the dependency of deviance
differences on scale and helps to pinpoint scale-independent or “robust” change-point locations. Thus mutagrams help in sorting out deviance differences that occur across many scales, pointing to actual changes, from those that are due to noise configurations on one hand or to smooth changes of the underlying function on the other hand. Mutagrams are related in spirit to mode trees (Minnotte and Scott, 1993) and mode forest (Minnotte, Marchette and Wegman, 1998) and also inspired by the Sizer idea of Chaudhuri and Marron (1999,2000).

These graphical ideas will be illustrated in the following with various sequence data, ranging from DNA sequences to coal mining disaster data.

2. ONE-SIDED QUASI-LIKELIHOOD

In this paper, deviance associated with local quasi-likelihood fitting plays a central role. A version of local quasi-likelihood was proposed in Fan, Heckman and Wand (1995); Staniswalis (1989) introduced the concept of a local likelihood. The quasi-likelihood function was first proposed by Wedderburn (1974), see also McCullagh and Nelder (1989), and is used in statistical inference when it is desirable to avoid full specification of a distribution family such as the exponential family, as required for likelihood and classical generalized linear models. The only specifications necessary for this single index model are link function $g(\cdot)$ and variance function $V(\cdot)$.

Given independent observations $(x_i, y_i)$ of predictors $x_i \in \mathbb{R}^p$, $p \geq 1$ and of responses $y_i \in \mathbb{R}$, $i = 1, \ldots, n$, the regression function is modeled as

$$\mu(x) = E(Y|X = x) = g(x^T\beta),$$

where $\beta = (\beta_0, \ldots, \beta_p)^T \in \mathbb{R}^{p+1}$ is a parameter vector and $x = (1, x^T)^T$ is the vector of predictor values, augmented by a 1. In this model, $\eta = x\beta$ plays the role of a linear predictor. Besides this specification of the conditional mean, the conditional variance is
assumed to depend solely on the conditional mean via $\text{Var}(Y|X = x) = \sigma^2 V(\mu(x))$, where $\sigma$ is a dispersion parameter.

Given link and variance functions, the quasi-likelihood for the observation $y_i$ is defined as

$$Q(\eta; y_i) = \tilde{Q}(\mu; y_i) = \int_{y_i}^{\mu} \frac{y_i - t}{\sigma^2 V(t)} \, dt,$$

and we define the quasi-deviance $\tilde{D} = -2\sigma^2 \tilde{Q}(\mu; y_i)$. The quasi-likelihood for the sample is then

$$\tilde{D}(\mu, y) = \sum_{i=1}^{n} \tilde{D}(\mu; y_i).$$

Localizing quasi-likelihood using local polynomial fitting, we obtain the local quasi-likelihood estimates

$$(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)(x) = \arg\max_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^{n} Q(\eta_i; y_i) w_i K(\frac{x_i - x}{h}),$$

where $\eta_i = \beta_0 + \beta_1 (X_i - x) + \ldots + \beta_p (X_i - x)^p$ (Braun and Müller, 1998). This simply means that the quasi-likelihood is applied only in a small window around the target location $x$. Here $K$ is a nonnegative kernel function for which we assume a compact support $[-1,1]$, and $h$ is a bandwidth defining the size of the window. The $w_i$ denote case weights which may be included in some cases, for example when the data are averages formed from unequal numbers of observations. The order of the local polynomial $p$ is usually chosen as $p = 0$ (locally constant) or $p = 1$ (locally linear). Then we set $\hat{\mu}(x) = g(\hat{\beta}_0(x))$, i.e., only the intercept $\hat{\beta}_0$ is needed.

We extend this to one-sided local quasi-likelihoods, where only data on the left or on the right of the target location enter. Similar ideas were considered in Braun and Müller (1998) and Loader (1999). Fixing a target argument $x$ and bandwidths $h$, the relevant windows are $[x - h, x]$ for the one-sided left quasi-likelihood and $[x, x + h]$ for the one-sided right quasi-likelihood. Formally, the one-sided quasi-likelihood estimates to the right and to the left are defined as
\[
(\hat{\beta}_{0+}, \hat{\beta}_{1+}, \ldots, \hat{\beta}_{p+})(x) = \arg\max_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^{n} 1_{\{x_i > x\}} Q(\eta_i; y_i) w_i K(\frac{x_i - x}{h})
\]

and

\[
(\hat{\beta}_{0-}, \hat{\beta}_{1-}, \ldots, \hat{\beta}_{p-})(x) = \arg\max_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^{n} 1_{\{x_i \leq x\}} Q(\eta_i; y_i) w_i K(\frac{x_i - x}{h}).
\]

We then define the maximizing values

\[
\hat{Q}(x, y; h) = \sum_{i=1}^{n} Q(\hat{\beta}_0(x) + \hat{\beta}_1(x)(x_i - x) + \ldots + \hat{\beta}_p(x)(x_i - x)^p; y_i) w_i K(\frac{x_i - x}{h})
\]

and

\[
\hat{Q}_\pm(x, y; h) = \sum_{i=1}^{n} 1_{\{x_i - x > 0\}} Q(\hat{\beta}_{0\pm}(x) + \hat{\beta}_{1\pm}(x)(x_i - x) + \ldots + \hat{\beta}_{p\pm}(x)(x_i - x)^p; y_i) w_i K(\frac{x_i - x}{h}),
\]

with corresponding fitted values, \(\hat{\mu}(x, h) = g(\hat{\beta}_0(x))\), \(\hat{\mu}_\pm(x, h) = g(\hat{\beta}_{0\pm}(x))\) and \(\hat{\mu}_i = \hat{\mu}_i(x, h), \hat{\mu}_i = \hat{\mu}_\pm(x, h)\).

The corresponding quasi-deviances

\[
D_\pm(x, y; h) = -2\sigma^2 \hat{Q}_\pm(x, y; h), \quad D(x, y; h) = -2\sigma^2 \hat{Q}(x, y; h)
\]

measure the deviance of the fitted model. The quasi-deviances do not depend on \(\sigma^2\). The most common quasi-deviances are (see McCullagh and Nelder, 1989) \(D(x, y; h) = \sum (y_i - \hat{\mu}_i)^2\) for normal data; \(D(x, y; h) = \sum -2(y_i \log \hat{\mu}_i - \hat{\mu}_i)\) for Poisson data; and \(D(x, y; h) = \sum (y_i \log(\frac{\hat{\mu}_i}{1 - \hat{\mu}_i}) + \log(1 - \hat{\mu}_i))\) for binomial data.

### 3. LOCAL CHANGE-POINT STATISTICS

In the following, we adopt the identity link and consider two local change-point approximations in order to define the local quasi-likelihood fits.
Local Jump Model. The polynomial that is fitted locally is a constant with $D(x, y; h) = \sum_{i=1}^{n} D(\hat{\beta}_0; y_i) w_i K(\frac{x_i-x}{h})$ and $\hat{\gamma} = \hat{\mu} = \hat{\beta}_0 = \bar{y}$. Fitting constants in the one-sided windows $[x-h, x]$ and $[x, x+h]$ will then allow to locally approximate a discontinuity at $x$, and the resulting fit will be compared with fitting a constant over the entire window, which is the fit without change-point.

Local Slope Change Model. A slope change is approximated by a function corresponding to a two-stick regression line with a change-point located at the midpoint $x$, given by $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 (x_{i1} - x) x_{i2}$, where $x_{i1} = x_i$ and $x_{i2} = 1_{(x_i > x)}$. The quasi-likelihood in this case equals $D^*(x, y; h) = \sum_{i=1}^{n} D(\beta_0 + \beta_1 x_{i1} + \beta_2 (x_{i1} - x) x_{i2}; y_i) w_i K(\frac{x_i-x}{h})$.

While the local jump model provides an approximation to a jump discontinuity, the local slope model serves as an approximation to a jump in the first derivative. Jumps in higher order derivatives could be modeled similarly but are of less interest. The slope change model reflects a slope change occurring at $x$, while the function itself remains continuous. This model fit will be compared with fitting an unbroken line over the entire window, a fit without change-point.

For each of these models, we fit three local quasi-likelihoods: Left-, right- and two-sided. An illustration of these fits is provided in Fig. 1 for the Nile data example (Cobb, 1978). The annual flow volume of the Nile river has been previously shown to have a change-point at the year 1898 by many authors. Given a chosen value for the bandwidth $h$ (we choose it within 1% to 40% of the range of the predictor variable), “evidence” functions $\Delta(x, h)$ providing evidence for the presence of a change-point at each location $x$ will be translated into change trees.

Choosing one of the above models, we obtain deviances $D_+(x, y; h)$, $D_-(x, y; h)$ and $D(x, y; h)$ of locally one- and two-sided model fits and for the Local Jump Model define the evidence function by
\[ \Delta(x, h) = |(D_+(x, y; h) + D_-(x, y; h)) - D(x, y; h)|. \]

For the local slope change model, we define analogously
\[ \Delta(x, h) = |D^*(x, y; h) - D_*(x, y; h)|, \]
where \( D^* \) is defined as above and is the deviance for the model with slope change, fitted to the data in the entire window, and \( D_* \) is the deviance for an unbroken line without slope change, also fitted to the data in the entire window,
\[ D_*(x, y; h) = \sum_{i=1}^{n} D(\beta_0 + \beta_1(x_i - x); y_i) w_i K\left(\frac{x_i - x}{h}\right). \]

Note that the function \( \Delta(\cdot, h) \) depends on the bandwidth \( h \) and quantifies the degree of evidence that separate left- and right-sided model fits are better than an overall local fit. The improvement is measured as the amount of decline in the deviance when allowing for the more flexible one-sided fits. In the case of fitting constants (Local Jump Model), we compare the deviances between a model that allows for separately fitted constants in left and right half-windows with a model that fits one global constant to the data in the entire window by maximum quasi-likelihood.

In practice, we obtain \( \Delta(x, h) \) on a grid of \( n \) points in the domain \([0, T]\) of the predictor variable, \( 0 < x_1 < x_2 < \ldots < x_n < T \), which is usually equidistant for sequence data. The maximization is carried out over this finite grid of points. Let for a small \( h_0 \ll h \)
\[ S_0 = [0, T] \setminus \{(0, h_0) \cup (T - h_0, T)\}, \]
and \( \tau_1 = \text{argmax}_{x \in S_0} \Delta(x, h) \). Then we adjust the area where possible change-points will be located to \( S_j = S_{j-1} \setminus (\tau_j - h_0, \tau_j + h_0) \), where \( \tau_j = \text{argmax}_{x \in S_{j-1}} \Delta(x, h) \) for \( j = 1, 2, \ldots, M \), and where the number of change-point locations that are being considered is \( M \) which will depend on \( h_0 \). This ensures a minimum distance between neighboring estimated change-points and also a minimum distance to the endpoints of the domain. If \( (\tau_1, \ldots, \tau_M) \) is the sequence of change-point locations identified in this way, given \( h \), the ordered evidence values \( \Delta(\tau_1, h), \ldots, \Delta(\tau_M, h) \) then form the basis for the construction of change trees.
4. CHANGE TREES

A change tree is a graph that provides a visual summary of the location of possible change-points in a data sequence. It is related in spirit to the mode tree (Minnotte and Scott, 1993; Marchette and Wegman, 1997). Given a point \( x \), the evidence that a change-point is located at \( x \) is based on the quantities \( \Delta(x, h) \) defined above. This evidence therefore depends on a bandwidth, on the choice of link function variance function and last not least, on the linear predictor function - in our case local constants, local two-stick regressions or local linear regressions.

Once \( (\tau_j, \Delta(\tau_j, h)), j = 1, 2, \ldots, M, \) have been determined, we construct vertical line segments defining the change tree. These are drawn from the location points \( \tau_j \) \( (j = 1, \ldots, M) \) to the associated values of observed change, \( \Delta(\tau_j, h) \). Horizontal segments join a vertical segment with the nearest vertical segment with a higher change value. That is, we join the vertical segment of \( \tau_j \) with that of \( \tau_{j+1} \) and \( \tau_{j+2} \). By doing this we construct a change tree that resembles the mode tree of Minnotte and Scott (1993).

The values of \( \Delta \) can be thresholded in order to “vertically” pr une the tree, in the sense that a location where \( \Delta(\tau_j, h) < \Delta_0 \) for a threshold \( \Delta_0 \) is discarded from further consideration. As a rule of thumb, we choose \( \Delta_0 = 0.1 \max_{1 \leq j \leq M} \Delta(\tau_j, h) \), and \( h_0 \) which effects “horizontal pruning” was chosen as \( h_0 = 0.04n \). For change trees based on the Local Jump Model, locations where \( \hat{\beta}_{0+} > \hat{\beta}_{0-} \), i.e., where changes from lower to larger level are observed, are indicated by a solid vertical line; locations where \( \hat{\beta}_{0+} < \hat{\beta}_{0-} \) are denoted by a dotted vertical line. For change trees based on the Local Slope Change Model, locations where the slope increases are indicated with a dashed line, and locations where it decreases, with a solid vertical line.

As an example, consider a biological sequence, the Bacteriophage \( \lambda \) Data \( (n = 485) \) from Skalka et al. (1968), see Fig. 2. The data correspond to Guanine and Cytosine (G+C) proportions which have been aggregated from the DNA sequence of this phage. Braun and Müller (1998) used split local polynomial fits for the Bacteriophage \( \lambda \) data and showed evidence of a change-point at 22.6Kbp and 33.2Kbp (kilo base pairs). A global step
function fitted with quasi-likelihood for these data was described in Braun et al. (2000).
The global segmentation methods are numerically tedious and local methods such as those
considered here provide for vastly increased flexibility and substantially faster computation.
Change trees for these data based on Local Jump Model (Fig. 3) and Local Slope Change
Model (Fig. 4) are constructed using the Binomial quasi-likelihood, with M=100 and
$h_0 = 1.9\text{Kbp}$. The changes observed here are in agreement with earlier findings; and both
change trees are seen to provide a useful visual summary for these changes. We also note
that Local Jump and Local Slope Change models show remarkable agreement for these
data.

5. MUTAGRAMS

Change trees depend on a bandwidth choice and by viewing one tree one cannot fully appreciate the effect that different bandwidths might have on sequence segmentation. A graphical approach to address this issue is to display a whole range of deviance changes simultaneously for different bandwidths at each support point $x$. The resulting graphical device is the Mutagram.

The mutagram is related in spirit to the mode forest (Minnotte, Marchette and Wegman, 1998). It is a graphical representation of a 2-dimensional matrix. For given coordinate points $x$ and $h$, define $M_{xh} = \Delta(x, h)$, with evidence function $\Delta(\cdot, \cdot)$ as defined in Section 3.
The matrix entries are based on a grid of bandwidths $h_i$ (in descending order), $i = 1, \ldots, m$, and locations $x_j, j = 1, \ldots, n$. The function $\Delta(x, h)$ is plotted as an image, where shades of gray represent the values of $\Delta$. The larger the value, the darker the shade.

Examples of mutagrams for the bacteriophage $\lambda$ data are provided in Fig. 5 for the Local Jump Model and in Fig. 6 for the local slope change model. We find that the same major change occurring at 22Kbp is identified in both models, and this value was also identified as major change-point in this sequence in by previous analyses. The Local Slope Change Model provides additional detail for changes that occur to the right of this major
change. The presence of such additional change-points is consistent with earlier analyses of these data (e.g., Braun et al., 2000), and indicates that the Local Slope Change Model is slightly better adapted to these data.

Mutagrams allow to simultaneously study the effect of location and bandwidth or scale, and follow the scaling philosophy of SiZer (Chaudhuri and Marron, 1999, Chaudhuri and Marron, 2000). Mutagrams pick up change-points as dark zones and surround them with a “shadow” of neighboring points that becomes more visible for larger bandwidths in the case of major change-points. This is due to the fact that a major change-point will exercise its impact more on larger windows than smaller windows, leading to increased impact in the mutagrams for larger bandwidths. Mutagrams are a useful and simple graphical tool to assess the nature of changes that occur in sequence data.

6. FURTHER ILLUSTRATIONS

6.1 Dow Jones Industrial Average
There exists a sizable literature on change-point models applied to stock market data. For example, Chen and Gupta (1997) used the Schwarz information criterion to detect change-points in stock market data, while Hsu (1979) and Lee (1998) detected a variance shift during the week of 19-23 March 1973.

We choose as sequence data the weekly closing of the Dow Jones Industrial Average from 1950 to 1999 (Fig. 7). The first step is the customary log transformation, applied to the weekly closing values ($n = 2609$ weeks) and then compute the local quasi-normal likelihoods to obtain the evidence function and change trees.

One finding is that, not surprisingly, the Local Slope Change Model is more suited to these data as compared to the Local Jump Model. The tree indicates that there are several major slope changes around 1965 and 1980. The effect of thresholding or pruning the tree can be seen by comparing Fig. 8, which has the full change tree, with Fig. 9, in which a pruned tree is displayed.
The mutagram based on the Local Slope Change Model (Fig. 10) reinforces the idea that there are slope changes around 1965 and 1980. These dates indicate the beginning and end of a phase between 1965 and 1980 where stocks were relatively flat.

6.2 S. Cerevisiae III

The S. Cerevisiae III \((n = 526)\) is a sequence of DNA of brewer’s yeast. The data we analyze reflects the Guanine and Cytosine \((G+C)\) proportion in the DNA sequence (Oliver et al., 1992) and are shown in Fig. 11. These data are binned and correspond to relative frequencies over the bins. The data set is available at the Genbank database (http://www.ncbi.nlm.nih.gov/Genbank/). DNA segmentation for these data has also been studied by Braun et al. (2000), Chechetkin and Lobzin (1998) and Liö et al. (1996). The Binomial quasi-likelihood is used in our analysis. In this example, the observed changes for both Local Jump Model (Fig. 12) as well as Local Slope Change Model (Fig. 13) are similar.

The mutagrams for the Local Jump Model and the Local Slope Change Model are presented in Fig. 14 and Fig. 15. Similar to the change-trees, the mutagrams also differ little between these two models, leading to relatively robust conclusions about the changes. Two well-defined change locations are detected from the mutagrams at 300 Kb and at about 420 Kb. Another change location may occur at 76 Kb but is less clearly defined.

6.3 Coal Mining Data

Maguire et al. (1952) reported the time intervals (in days) between successive British coal mining accidents with 10 or more casualties. Jarrett (1979) extended these data to include all accidents between March 1851 and March 1962. From the scatter plot (Fig. 16) it appears that there are several periods of sharp changes in the frequency of accidents.

Different approaches have been used for analyzing this data set. Graphically based interval estimation (Boukai, 1994), Bayesian analysis (Lee, 1998) and conditional and unconditional maximum likelihood (Jandhyala et al. 2000) all have identified 1898 as a major change-point. Here we obviously have count data and therefore applied the Poisson quasi-
likelihood. The Local Jump Model seemed to be most appropriate here, leading to the change tree as shown in Fig. 17 and the mutagram of Fig. 18. Both graphical tools point to two periods where rapid changes occur, around the years 1895 and 1950.

7. CONCLUSIONS

Change trees and mutagrams are simple graphical tools that nevertheless serve to enhance the detection, description and visualization of changes in sequence data. Such data originate from diverse fields. Since the local likelihoods are determined based on quasi-deviance, the method adapts to data with non-normal distributions. The observed change features depend mainly on the threshold that is employed to declare a change and on the window width. These graphical tools allow to gauge the influence of these auxiliary parameters and to deduce characteristics of the changes, similar to the mode tree of Minnotte and Scott (1993).

In particular, mutagrams allow to gauge the stability of a “change feature” across different scales, taking into account a whole range of window widths simultaneously. Regarding a comparison between the two local change models we are considering, namely the Local Jump Model and the Local Slope Change model, we find that they are equally effective in some applications (S.Cerevisiae), while for other sequences the Slope Change Model is clearly more appropriate, for example when considering the Dow Jones weekly closing data.

On the other hand, the Local Jump Model is also important in some situations such as for the coal mining disaster data. We conclude that both models are useful and complement each other. Questions that might be addressed in future investigations are the asymptotic properties of the local change-point detection method based on the differences in quasi-likelihood fits and the extension to models and likelihoods which allow the explicit incorporation of dependence structure that may be often present in observed sequences is of interest.
Fig. 1. The Local Jump Model (above) and the Local Slope Change Model (below), displaying the relevant fits for the Nile data (Cobb 1978), centered at the year 1898 with bandwidth 10 years.

Fig 2. The G+C proportions of Bacteriophage λ, as a function of genome location, measured in kilo base pairs (Kbp) (Skalka et al., 1968).
Fig 3. Change Tree for the G+C proportion in the DNA sequence of Bacteriophage λ, using the Local Jump Model with bandwidth=5 Kbp.

Fig 4. Change Tree for Bacteriophage λ, using the Local Slope Change Model with bandwidth=5 Kbp.
Fig 5. Mutagram for the Bacteriophage lambda DNA sequence using the Local Jump Model. Bandwidths range from 5-20Kbp (10% to 40% of the data range).

Fig 6. Mutagram for the Bacteriophage lambda DNA sequence using the Local Slope Change Model (bandwidths as in Fig. 5).

Fig 8. Change Tree for Dow Jones Industrial Average, using the Local Slope Change Model with bandwidth=10 years.
Fig 9. Pruned Change Tree for Dow Jones Industrial Average, using the Local Slope Change Model with bandwidth=10 years.

Fig 10. Mutagram for the Dow Jones Industrial Average using the Local Slope Change Model.
Fig 11. Scatter plot for the G+C proportion of the DNA sequence of S. Cerevisiae III (Oliver et al., 1992).

Fig 12. Change Tree for the S. Cerevisiae III sequence, using the Local Jump Model with bandwidth=50 Kbp.
Fig 13. Change Tree for the S. Cerevisiae III sequence, using the Local Slope Change Model with bandwidth=50 Kbp.

Fig 14. Mutagram for the S. Cerevisiae III sequence, using the Local Jump Model.
Fig 15. Mutagram for the S. Cerevisiae III sequence, using the Local Slope Change Model.

Fig 16. Scatter plot for the number of coal mining accidents in Britain, 1850-1962 (Maguire et al., 1952).
Fig 17. Change Tree for number of coal mining accidents, using the Local Jump Model with bandwidth=20 years.

Fig 18. Mutagram for the number of coal mining accidents, using the Local Jump Model.
REFERENCES


