Methods and algorithms for robust filtering

Ursula Gather

Department of Statistics, University of Dortmund, 44221 Dortmund, Germany

Roland Fried

Department of Statistics, University Carlos III, 28903 Getafe (Madrid), Spain

Abstract: We discuss filtering procedures for robust extraction of a signal from noisy time series. Moving averages and running medians are standard methods for this, but they have shortcomings when large spikes (outliers) respectively trends occur. Modified trimmed means and linear median hybrid filters combine advantages of both approaches, but they do not completely overcome the difficulties. Improvements can be achieved by using robust regression methods, which work even in real time because of increased computational power and faster algorithms. Extending recent work we present filters for robust online signal extraction and discuss their merits for preserving trends, abrupt shifts and extremes and for the removal of spikes.

Key words: Signal extraction, drift, edge, outlier, update algorithm

1 Introduction

In speech recognition, video transmission and intensive care monitoring the basic task is to extract a signal from the observed noisy time series. The signal is assumed to vary smoothly most of the time with a few abrupt shifts. Besides the attenuation of normal observational noise and the removal of outlying spikes for recovering smooth sequences, the preservation of the locations and heights of shifts and local extremes is important. All this needs to be done automatically and in real time with short delays. This increases the risk of confusing outlier sequences and shifts or local extremes. For distinguishing extremes and outliers we rely on the smoothness of the underlying signal, i.e. observations which are far away from an estimated signal value are treated as outliers and not as being due to a signal peak. We can identify shifts by their duration setting a lower limit for the length of a relevant shift.

Moving averages and other linear filters are popular for signal extraction as they recover trends and are very efficient in Gaussian samples, but they are highly vulnerable to outliers and they blur level shifts. Tukey (1977) suggests running medians for removing outliers and preserving level shifts, but standard medians have deficiencies in trend periods (Fried and Gather, 2002). Linear median hybrid filters (Heinonen and Neuvo, 1987, 1988) have been suggested as they are computationally more efficient than running medians, and preserve shifts similarly good or even better than these. These filters track polynomial trends, but they can only remove single isolated outliers. Modified trimmed mean filters are another compromise between running means and running medians. They choose an adaptive amount of trimming, but like running medians they also deteriorate in trend periods.

A better solution for tracking trends is to replace the median, a robust location estimator, by the estimated intercept obtained by robust regression of the data in a moving window against time. Based on a comparison of functionals with high breakdown point Davies, Fried and Gather (2004) recommend Siegel's (1982) repeated median because of its robustness against outliers and its stability. Since larger outliers have stronger effects on the repeated median we can add automatic rules for online trimming of outliers and construct procedures which are almost as bias-robust as filters based on least median of squares regression

(Rousseuw, 1984), but considerably faster and more efficient for Gaussian samples (Fried, 2004). The Q_α-method (Croux and Rousseeuw, 1992, Rousseeuw and Croux, 1993) has very nice properties for scale estimation even when a level shift occurs (Gather and Fried, 2003).

Robust regression also allows to construct hybrid filters which have similar benefits as linear median hybrid filters, while being considerably more robust (Fried, Bernholt and Gather, 2004). Procedures applying adaptive trimming which do not deteriorate in trend periods can also be derived (Bernholt, Fried, Gather and Wegener, 2004b).

In Section 2 we introduce the filtering procedures. In Section 3 we discuss computational and other aspects. In Section 4 we propose a robust rule for the adaptive choice of the window widths. In Section 5 we analyze real and simulated data for further comparison before we give some conclusions.

2 Methods for robust filtering

We assume a component model for the sequence (x_t) of observed data

$$x_t = \mu_t + u_t + v_t,\ t \in \mathbb{Z}. \tag{1}$$

The underlying signal μ_t is the level of the time series, which is assumed to vary smoothly with a few sudden changes, while u_t is additive noise from a symmetric distribution with mean zero and variance σ^2, and v_t is impulsive (spiky) noise from an outlier generating mechanism. For online signal extraction we move a time window of width $n = 2k + 1$ through the series and use $x_{t-k}, \ldots, x_{t+k}$ to approximate μ_t. This causes a time delay of k observations. Firstly we fix k to a given value for all filters.

2.1 Filters based on robust regression

A standard median filter (running median) approximates the signal μ_t by the median of the observations $\{x_{t-k}, \ldots, x_{t+k}\}$ within a moving time window,

$$SM(x_t) = \tilde{\mu}_t = med\{x_{t-k}, \ldots, x_{t+k}\},\ t \in \mathbb{Z},$$

where μ_t is regarded as the level of the series at time point t, which is assumed to be locally constant. For tracking trends, Davies et al. (2004) suggest fitting a local linear trend $\mu_{t+i} = \mu_t + i\beta_t,\ i = -k, \ldots, k$, to $\{x_{t-k}, \ldots, x_{t+k}\}$ by robust regression and recommend Siegel's (1982) repeated median (RM). When applied to the data (i, x_{t+i}), $i = -k, \ldots, k$, the RM reads

$$\begin{aligned} RM(x_t) = \tilde{\mu}_t^{RM} &= med\{x_{t-k} + k\tilde{\beta}_t, \ldots, x_{t+k} - k\tilde{\beta}_t\} \\ \tilde{\beta}_t^{RM} &= med_{i=-k,\ldots,k}\left\{med_{j\neq i}\frac{x_{t+i} - x_{t+j}}{i-j}\right\}. \end{aligned}$$

2.2 Filters based on trimming

Lee and Kassam (1985) suggest modified trimmed mean (MTM) filtering as a compromise between running means and running medians. MTM filters regulate the amount of trimming depending on the data. Firstly the local median $\tilde{\mu}_t$ and the local median absolute deviation about the median (MAD) $\tilde{\sigma}_t$ are calculated, then all observations farther away from the median than a multiple $q_t = d\tilde{\sigma}_t$ of the MAD are trimmed. Finally, the average of the remaining observations is taken as filter output:

$$
\begin{aligned}
MTM(x_t) &= \frac{1}{n_t}\sum_{i=-k}^{k} x_{t+i}\cdot 1_{[\tilde{\mu}_t-q_t,\tilde{\mu}_t+q_t]}(x_{t+i}),\\
n_t &= \#\{x_{t+i}\in[\tilde{\mu}_t-q_t,\tilde{\mu}_t+q_t], i=-k,\dots,k\},\\
q_t &= d\cdot c_n\cdot med\{|x_{t-k}-\tilde{\mu}_t|,\dots,|x_{t+k}-\tilde{\mu}_t|\}\,.
\end{aligned}
$$

Here, c_n is a correction factor, which is chosen to achieve unbiasedness for Gaussian noise. For a very large window width we get $c_n = 1.483$, while e.g. for $n = 21$ we have $c_n = 1.625$. For $d = 0$, $MTM(x_t)$ is a running median, while for $d = \infty$ we get a moving average.

MTM filters implicitly assume a location model as standard median filters do. A straightforward modification is to fit a local linear trend by the repeated median and trim those observations having large residuals in this regression setting. The local variability can be estimated by applying the MAD to the regression residuals (Bernholt et al., 2004). The filter output can then be derived either by least squares regression or by the repeated median of the observations with moderately large residuals. We denote the resulting filters by TRM and MRM, respectively:

$$
\begin{aligned}
TRM(x_t) &= \overline{x}_{J_t} - \tilde{\beta}_t^{TRM}\overline{j}_{J_t}\\
\overline{x}_{J_t} &= \frac{1}{|J_t|}\sum_{j\in J_t} x_{t+j}\\
\overline{j}_{J_t} &= \frac{1}{|J_t|}\sum_{j\in J_t} j\\
\tilde{\beta}_t^{TRM} &= \frac{\sum\limits_{j\in J_t}(j-\overline{j}_{J_t})(x_{t+j}-\overline{x}_{J_t})}{\sum\limits_{j\in J_t}(j-\overline{j}_{J_t})^2}\\
J_t &= \{j=-k,\dots,k : |x_{t+j}-\tilde{\mu}_t^{RM}-j\tilde{\beta}_t^{RM}|\le q_t\}\\
MRM(x_t) &= med\{x_{t+j}-j\tilde{\beta}_t^{MRM},\ j\in J_t\}\\
\tilde{\beta}_t^{MRM} &= med_{i\in J_t}\left\{med_{j\in J_t, j\neq i}\frac{x_{t+i}-x_{t+j}}{i-j}\right\},
\end{aligned}
$$

with $(\tilde{\mu}_t^{RM}, \tilde{\beta}_t^{RM})$ being the repeated median level and slope estimate for the current time window $\{x_{t-k},\dots,x_{t+k}\}$.

2.3 Hybrid filters

Linear median hybrid filters are combinations of linear and median filters (Heinonen and Neuvo, 1987, 1988). Linear subfilters are applied to the input data before taking the median of their outcomes as final filter output. This reduces computation time and increases the flexibility compared to standard median filters due to the variety of linear subfilters. Linear median hybrid filters with finite impulse response, briefly FMH filters, are characterized by subfilters which respond to a finite number of impulses only.

A simple FMH filter corresponds to a location model and applies two one-sided moving averages and the current observation x_t as central subfilter for edge preservation:

$$SFMH(x_t) = med\{\Phi_1(x_t),\ x_t,\ \Phi_2(x_t)\}$$

$$\Phi_1(x_t) = \frac{1}{k}\sum_{i=1}^{k} x_{t-i}, \qquad \Phi_2(x_t) = \frac{1}{k}\sum_{i=1}^{k} x_{t+i}.$$

Predictive FMH filters correspond to a linear trend model and apply predictive FIR subfilters for one-sided extrapolation of a trend:

$$PFMH(x_t) = med\{\Phi_F(x_t), x_t, \Phi_B(x_t)\},$$
$$\Phi_F(x_t) = \sum_{i=1}^{k} h_i x_{t-i}, \qquad \Phi_B(x_t) = \sum_{i=1}^{k} h_i x_{t+i}.$$

Choosing the weights $h_i = \frac{4k-6i+2}{k(k-1)}$, $i = 1, \ldots, k$ results in the minimal mean square error (MSE) predictions for a linear trend which is disturbed by white noise (Heinonen and Neuvo, 1988).
Combined FMH filters use predictions of different degrees,

$$CFMH(x_t) = med\{\Phi_F(x_t), \Phi_1(x_t), x_t, \Phi_2(x_t), \Phi_B(x_t)\},$$

with $\Phi_1(x_t)$, $\Phi_2(x_t)$, $\Phi_F(x_t)$ and $\Phi_B(x_t)$ being the subfilters for forward and backward extrapolation of a constant signal and a linear trend as given above.

In view of increased computational power and because of improved algorithms, computation time is nowadays not a great problem. Fried, Bernholt and Gather (2004) use half-window medians and repeated medians to construct robust hybrid filters:

$$\begin{aligned} PRMH(x_t) &= med\{RM^F(x_t), x_t, RM^B(x_t)\} \\ CRMH(x_t) &= med\{RM^F(x_t), \tilde{\mu}_t^F, x_t, \tilde{\mu}_t^B, RM^B(x_t)\} \end{aligned}$$

Here, $\tilde{\mu}_t^F = med\{x_{t-k}, \ldots, x_{t-1}\}$ and $\tilde{\mu}_t^B = med\{x_{t+1}, \ldots, x_{t+k}\}$ are half-window medians, while $RM^F(x_t)$ and $RM^B(x_t)$ estimate the level at time t using the repeated median of $x_{t-k}, \ldots, x_{t-1}$ and $x_{t+1}, \ldots, x_{t+k}$, respectively:

$$\begin{aligned} RM^F(x_t) &= med\{x_{t-k} + k\tilde{\beta}_t^F, \ldots, x_{t-1} + \tilde{\beta}_t^F\} \\ \tilde{\beta}_t^F &= med_{i=-k,\ldots,-1}\left\{med_{j=-k,\ldots,-1,j\neq i}\frac{x_{t+i}-x_{t+j}}{i-j}\right\} \\ RM^B(x_t) &= med\{x_{t+1} - \tilde{\beta}_t^B, \ldots, x_{t+k} - k\tilde{\beta}_t^B\} \\ \tilde{\beta}_t^B &= med_{i=1,\ldots,k}\left\{med_{j=1,\ldots,k,j\neq i}\frac{x_{t+i}-x_{t+j}}{i-j}\right\}. \end{aligned}$$

3 Comparison of different filtering procedures

In the following we compare the previous filtering procedures w.r.t. computation time and their analytical properties.

3.1 Computation

The time needed for the filtering is crucial in real time applications. Fast algorithms for the update of the filter output are needed for online signal extraction. Denoting the length of the time window by n, the median of the proceeding window can be updated in logarithmic time ($O(\log n)$) using linear space if the data in the window are stored in sorted order using a red-black tree (Cormen, Leiserson and Rivest, 1990, Section 15.1). This improves on the linear time needed for calculating the median from scratch.

An algorithm for the update of the repeated median in linear time using quadratic space based on a hammock graph is proposed by Bernholt and Fried (2003), and another update algorithm needing only linear space running in $O(n \log n)$ average time is presented by Fried, Bernholt and Gather (2004).

Updating the MAD can be done in logarithmic time (Bernholt et al., 2004). Hence, the MTM can be calculated in logarithmic and the TRM in linear time. For the MRM, however, $O(n \log n)$ time is needed at least for the second repeated median. Detailed descriptions of the update algorithms can be found in Bernholt et al. (2004) and Fried, Bernholt and Gather (2004).

Table 1 summarizes the time and the space needed for the updates of the filtering procedures. Note that the space for the repeated median and both repeated median hybrid filters can be reduced to $O(n)$, but at the expense of larger computation times.

Table 1: Time and space needed for the update of the filters.

	SM	RM	MTM	MRM	TRM	FMH	RMH
time	$O(\log n)$	$O(n)$	$O(\log n)$	$O(n \log n)$	$O(n)$	$O(1)$	$O(n)$
space	$O(n)$	$O(n^2)$	$O(n)$	$O(n)$	$O(n^2)$	$O(n)$	$O(n^2)$

3.2 Analytical properties

For a discussion of the filtering procedures we concentrate on the analytical properties within a single time window when being applied to data generated from the component model (1).

Equivariance and invariance are important properties of statistical procedures. Location equivariance means that adding a constant to all observations in a window changes the filter output accordingly. Scale equivariance means that multiplication of all observations with a constant changes the estimate in the same way. All the above procedures possess these two properties.

Only some of the procedures are trend invariant, however (Fried, Bernhold and Gather, 2004a,b). This property means that the extracted level does not change when adding a linear trend as long as the central level is fixed. The RM, the PFMH, the PRMH, the TRM and the MRM are trend invariant, while the median, the MTM, the CFMH and the CRMH are not. Therefore, for the latter methods the efficiency, the removal of spikes and the preservation of shifts are influenced by underlying trends.

Filters which are not trend invariant blur e.g. upward shifts within downward trends. Although the median and the MTM can remove k spikes completely in a single time window from a constant signal if there is no observational noise ($\sigma^2 = 0$), even a single positive outlier within a downward trend causes smearing. The predictive FMH can remove a single spike and preserve a shift exactly within a linear trend irrespectively of the directions as it is trend invariant, while the combined FMH does so only if the outlier (shift) has the same direction as the trend. The RMH filters improve on the FMH filters as they can remove up to $\lfloor k/2 \rfloor$ subsequent spikes without any effect. Furthermore, the predictive RMH preserves shifts exactly, while for the combined RMH this is true only if the shift is in the same direction as the trend, just like for the combined FMH. The RM, the TRM and the MRM can even remove $k-1$ spikes completely within a single time window irrespectively of a linear trend if $\sigma^2 = 0$.

The previous results hold when there is no observational noise. Lipschitz continuity restricts the influence of minor changes in the data due to small noise or rounding. The standard median, the FMH, the RM and the RMH filters are Lipschitz-continuous. The median is Lipschitz-continuous with constant 1 like all order statistics, while the repeated median and the repeated median hybrid filters are Lipschitz-continuous with constant $2k+1$. An FMH filter is Lipschitz-continuous with constant $max|h_i^j|$, the maximal absolute weight given by a subfilter. MTM, MRM and TRM filters, however, are not Lipschitz-continuous, which can cause instabilities when there are small changes in the data. The discontinuity is caused by the trimming of observations. Application of continuous M-estimators is preferable for this reason, but computationally more expensive. Nevertheless, we investigate simpler trimming based methods in order to obtain information about the possible gain by further iterations.

The finite-sample breakdown point (FSBP) is the fraction of observations which have to be put into worst case positions in order to make the estimate take arbitrarily wrong values. For the median the breakdown point becomes $(k+1)/n$ when applied to $n = 2k+1$ data points, meaning that at least half of the window needs to be outlying in order to cause an arbitrarily large spike in the extracted signal. Since for the explosion of the local MAD also at least $k+1$ observations need to be modified, the MTM has the same breakdown point, while for the FMH filters two outliers are sufficient to make it break down. From the following Table we see that the RMH filters are considerably more robust than the FMH filters, and that the RM, TRM and MRM are almost as robust as the median in the sense of breakdown.

Table 2: Fraction of outliers in a window causing breakdown.

SM	MTM	RM	TRM	MRM	FMH	PRMH	CRMH
$\frac{k+1}{n}$	$\frac{k+1}{n}$	$\frac{k}{n}$	$\frac{k}{n}$	$\frac{k}{n}$	$\frac{2}{n}$	$\frac{\lfloor k/2 \rfloor + 1}{n}$	$\frac{\lfloor k/2 \rfloor + 2}{n}$

Simulations show the effect of the second step in the derivation of the TRM and the MRM on their MSE as compared to that of the RM filter. Application of least squares to the trimmed observations (TRM) increases the efficiency for Gaussian noise, but almost preserves the robustness of the repeated median, while application of the repeated median (MRM) further reduces the bias caused by outliers (Fried, Bernholt and Gather, 2004b).

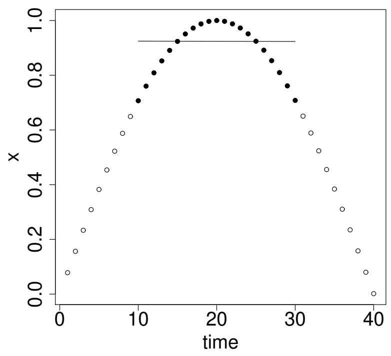
4 Adaptive choice of the window width

From the previous discussion we see that only the repeated median and the predictive hybrid filters PFMH and PRMH are both trend invariant and continuous, i.e. stable w.r.t. the occurrence of both trends and small changes in the data. The hybrid filters tend to preserve shifts and extremes, while the repeated median smoothes them considerably when being applied with a large window width (Fried, Bernholt and Gather, 2004a, b). This means that on the one hand we should choose a short window width, but on the other hand a large window width is better for removing outlier patches and for the attenuation of the observational noise. This is a robust variant of the common problem of bandwidth selection in nonparametric smoothing.

Fried (2004) investigates rules for online shift detection based on the most recent residuals in the time window. Similarly, we can formulate rules for the automatic choice of the window width using the regression residuals. Often least squares criteria are used to assess the local model fit and to find the bandwidth, but this is not suitable when outliers are present. Instead, a robust criterion is needed. Remembering that the median is the value which balances the signs of the residuals and that the repeated median is a regression analogue, it is natural to use the sign of the residuals. In this way we give the same weight to all observations irrespectively of their magnitude. However, note that there are always as many positive as negative residuals in the window for the repeated median fit. Therefore, we have to apply this idea to a suitable subset.

The Figure below visualizes the smoothing of a maximum by fitting a line with a too large window width. The residuals in the center will typically be positive, while most of the residuals at the start and the end of the window will be negative. These signs are simply al reverse for a minimum. Therefore it is natural to use the total number of positive residuals at the start and the end of the window for assessing the model fit. We divide the window into three sections as follows, namely the first $\lfloor (k+1)/2 \rfloor$ observations, the central $n - 2\lfloor (k+1)/2 \rfloor$ observations and the last $\lfloor (k+1)/2 \rfloor$ observations. If the total number T of positive residuals in the first and the last section is much larger than the average $\lfloor (k+1)/2 \rfloor$, we should shorten the window width since the signal slope might be decreasing substantially within the window. If T is much smaller than $\lfloor (k+1)/2 \rfloor$, the window width should also be shortened since the signal slope might be increasing.

Figure 1: Smoothing of a maximum by fitting a line to the filled points.



However, this reduction should not result in a window width which is too small to resist outlying patterns. Results of previous studies (Fried, Bernholt and Gather, 2004a, b) show that the repeated median resists up to between 25% and 30% outliers without being substantially affected. Therefore, the minimal window width should be about four times the maximal length of outlier patches to be removed. For patches of length three e.g. we use the constraint $n \geq 11$. Since longer time windows allow better attenuation of observational noise and also robustness against many outliers we increase the window width after each step whenever possible.

The proposed repeated median algorithm with robust adaptive selection of the window width is as follows: Let $k_l < k_u$ be lower and upper bounds for k, and $0 \leq d_l < 1 < d_u \leq 2$ be constants. Set $k = k_l$ and $t = k + 1$.

1. Calculate the repeated median fit $(\tilde{\mu}_t, \tilde{\beta}_t)$ for $x_{t-k}, \ldots, x_{t+k}$ to obtain $RM(x_t) = \tilde{\mu}_t$.

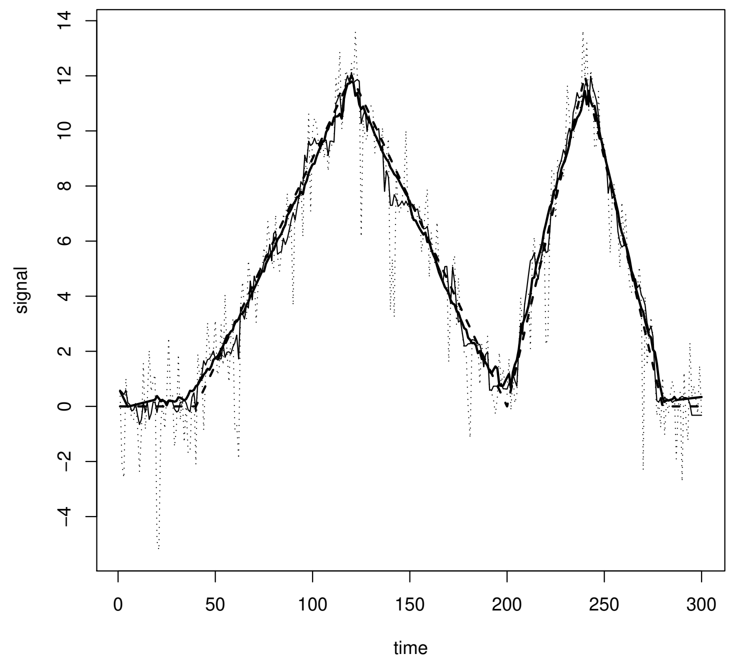
2. Get the residuals $r_i = x_{t+i} - \tilde{\mu}_t - i\tilde{\beta}_t$, $i = -k \ldots, k$, and set $T = \#\{i = -k, \ldots, -k - 1 + \lfloor (k+1)/2 \rfloor, k + 1 - \lfloor (k+1)/2 \rfloor, \ldots, k : r_i > 0\}$.
3. If $k > k_l$ and $T < d_l \cdot \lfloor (k+1)/2 \rfloor$ or $T > d_u \cdot \lfloor (k+1)/2 \rfloor$ set $k = k - 1$ and go to 1.
4. If $k < k_u$ set $k = k + 1$.
5. Set $t = t + 1$ and go to 1.

The same or similar approaches can be used for the other robust filters. We just need to modify the window sections for the hybrid filters possibly obtaining asymmetric filters.

5 Application

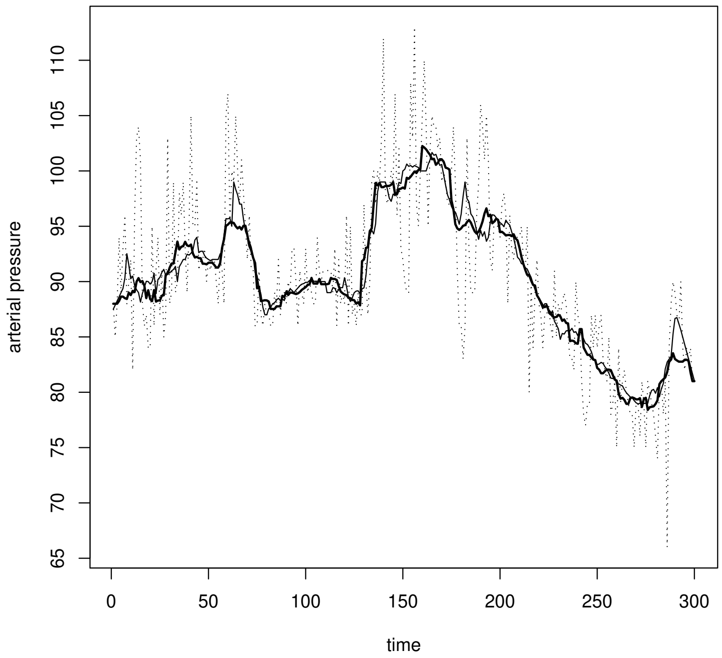
We now apply the filtering procedures to two data sets. The first example is a time series simulated from an underlying sawtooth signal, which is overlaid by Gaussian white noise with zero mean and unit variance, and there are three isolated, three pairs and two triples of outliers of size -5. The Figure below shows the outputs of the CRMH and the adaptive RM filter with $k_l = 5$, $k_u = 15$, $d_l = 0.7$ and $d_u = 1.3$. The CRMH with $n = 21$ preserves the local extremes very well, but it is rather variable. The adaptive RM is almost as good at the extremes while being much smoother. Most of the time a width close to the maximal $n = 31$ is chosen, but close to the three local extremes and at about t=280 the width decreases even to the minimal $n = 11$. The PRMH not shown here is similar to the CRMH, but it is more affected by the outliers, while the ordinary RM and the median cut the extremes.

Figure 2: Simulated time series (dotted), underlying signal (dashed) and outputs of the CRMH (thin solid) and the RM with adaptive window width (bold solid).



As a second example we analyze five hours of measurement of the arterial blood pressure of an intensive care patient. Figure 3 visualizes these data along with the outcomes of the MRM with a window width of $n = 21$ and of the adaptive RM filter with the same constants as before. The MRM resists some aberrant patterns very well, but it oversmoothes the local extremes at $t = 70$ and at $t = 290$. The adaptive RM again chooses the largest width $n = 31$ most of the time, but the width drops down to $n = 17$ about t=175 and t=225, and even to the minimal $n = 11$ about t=60 and t=130. It performs better at the extremes than the MRM, but it is affected by two subsequent outlying patterns about t=180. The RM with fixed window width also shows a spike there and performs in between the adaptive RM and the MRM at the extremes.

Figure 3: Arterial blood pressure (dotted) and outputs of the MRM (bold solid) and the RM with adaptive window width (thin solid).



6 Conclusion

Improved numerical algorithms render the real time application of robust procedures for time series filtering possible. Methods for robust regression like the repeated median allow to construct filters which have similar benefits like classical linear or location based approaches when these perform well, but overcome deficiencies w.r.t. the removal of spiky noise (outliers) or the tracking of trends. We find the repeated median procedure with robust adaptive choice of the window width particularly promising. First applications show that this algorithm can be modified even for online filtering without any time delay by estimating the intercept at the right hand side of the time window, but more experience is needed to optimize the automatic choice of the window width then.

Acknowledgement:
The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of complexity in multivariate data structures") is gratefully acknowledged.

References

[1] Bernholt, T., and Fried, R. (2003). *Computing the update of the repeated median regression line in linear time.* Information Processing Letters **88**, 111 – 117.

[2] Bernholt, T., Fried, R., Gather, U. and Wegener, I. (2004). *Modified repeated median filters.* Preprint, Department of Statistics, University of Dortmund, Germany.

[3] Cormen, T.H., Leiserson, C.E., and Rivest, R.L. (1990). *Introduction to Algorithms.* MIT Press, Cambridge, Massachusetts, and McGraw-Hill Book Company, New York.

[4] Croux, C., and Rousseeuw, P.J. (1992). *Time-efficient algorithms for two highly robust estimators of scale.* COMPSTAT 1992, Physica-Verlag, Heidelberg, 411 – 428.

[5] Davies, P.L., Fried, R., and Gather, U. (2004). *Robust signal extraction for on-line monitoring data.* J. Statistical Planning and Inference **122**, 65 – 78.

[6] Fried, R. (2004). *Robust filtering of time series with trends.* Current Advances and Trends in Nonparametric Statistics, special issue of the J. of Nonparametric Statistics, to appear.

[7] Fried, R., Bernholt, T., and Gather, U. (2004). *Repeated median and hybrid filters.* Computational Statistics & Data Analysis, to appear.

[8] Fried, R., and Gather, U. (2002). *Fast and robust filtering of time series with trends.* COMPSTAT 2002, Physica-Verlag, Heidelberg, 367 – 372.

[9] Gather, U., and Fried, R. (2003). *Robust estimation of scale for local linear temporal trends.* Proceedings of PROBASTAT 2002, Tatra Mountains Mathematical Publications **26**, 87 – 101.

[10] Heinonen, P., and Neuvo, Y. (1987). *FIR-median hybrid filters.* IEEE Transactions on Acoustics, Speech, and Signal Processing **35**, 832 – 838.

[11] Heinonen, P., and Neuvo, Y. (1988). *FIR-median hybrid filters with predictive FIR substructures.* IEEE Transactions on Acoustics, Speech, and Signal Processing **36**, 892 – 899.

[12] Lee, Y., and Kassam, S. (1985). *Generalized median filtering and related nonlinear filtering techniques.* IEEE Transactions on Acoustics, Speech, and Signal Processing **33**, 672 – 683.

[13] Rousseeuw, P.J. (1984). *Least median of squares regression.* J. American Statistical Association **79**, 871 – 880.

[14] Rousseuw, P.J., and Croux, C. (1993). *Alternatives to the median absolute deviation.* J. American Statistical Association **88**, 1273 – 1283.

[15] Siegel, A.F. (1982). *Robust regression using repeated medians.* Biometrika **68**, 242 – 244.

[16] Tukey, J.W. (1977). *Exploratory Data Analysis.* Addison-Wesley, Reading, Mass. (preliminary ed. 1971).