

# The Class of Generalized Hampel Filters

Ronald K. Pearson  
DataRobot  
Boston, MA, USA

Yrjö Neuvo  
Aalto University  
Espoo, Finland

Jaakko Astola, Moncef Gabbouj  
Tampere University of Technology  
Tampere, Finland

## Abstract

The standard median filter has only one tuning parameter—the width of the moving window on which it is based—and this has led to the development of a number of extremely useful extensions, including the recursive median filter, weighted median filters, and recursive weighted median filters. The Hampel filter is a member of the class of decision filters that, as we note here, may be viewed as another generalization of the median filter. This paper exploits this relationship, defining and briefly exploring the class of *generalized Hampel filters*, obtained by applying the median filter extensions listed above.

## 1 Introduction

In their paper, “On a class of nonlinear filters,” Sicuranza and Carini begin by noting [6]:

“The set of nonlinear filters is extremely large since their definition simply excludes the applicability of the linear superposition property on which the theory of linear filters is based. However, from the very beginning, attempts have been done to suitably classify nonlinear filters on the basis of some peculiar properties, leading to the identification of certain classes of nonlinear filters.”

This paper adopts a similar philosophy, restricting consideration to a class of nonlinear filters obtained by combining two previously studied filter classes: the Hampel filter described in Sec. 2, and the median filter extensions described in Sec. 3. The result is a

class of nonlinear filters we believe to be new, including these previously studied filters as special cases, but exhibiting a greater degree of design flexibility.

## 2 The Hampel filter

The Hampel filter belongs to the class of *decision-based filters* discussed in the book by Astola and Kosmanen [1, p. 194], who note that the basic concept has been reinvented again and again. The specific filter considered here is a moving-window implementation of the *Hampel identifier* described by Davies and Gather [2]: the central point in the data window  $\mathbf{w}_k$  is deemed an outlier and replaced if it lies more than  $t$  times the moving data window’s MAD scale estimate from its median:

$$y_k = \begin{cases} x_k & \text{if } |x_k - m_k| \leq tS_k. \\ m_k & \text{otherwise,} \end{cases} \quad (1)$$

Here,  $m_k$  is the median of  $\mathbf{w}_k$  and  $S_k$  is the MAD scale estimate, defined as:

$$S_k = 1.4826 \times \text{median}_{j \in [-K, K]} \{|x_{k-j} - m_k|\}, \quad (2)$$

Note that  $m_k$  is the output of the standard median filter, so the Hampel filter reduces to the standard median filter when  $t = 0$ .

To provide a basis for comparing the different filters considered in this paper, we apply them to the 420-point simulated data sequence shown in Fig. 1, which contains four components: a piecewise-linear “step-and-ramp” signal shown as the dark solid line in the plot, zero-mean Gaussian white noise with standard deviation  $\sigma = 0.1$  contaminating the first 240 points of the sequence, a sinusoidal component

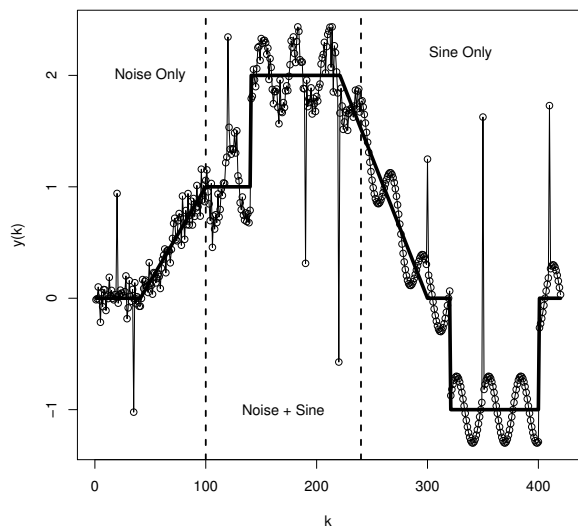
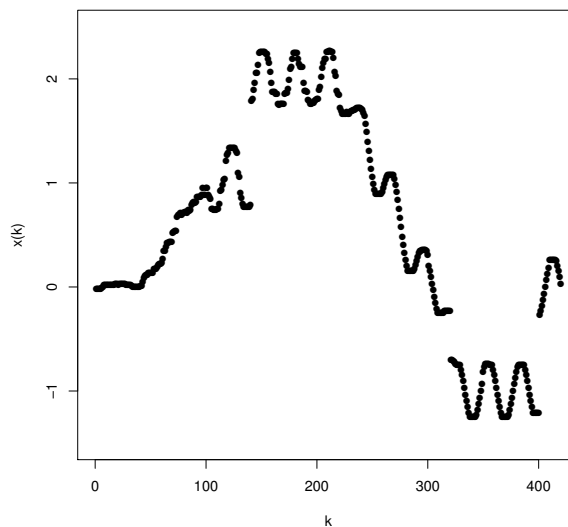


Figure 1: A simulated 420-point signal.

with period 29 and amplitude 0.3, appearing from  $k = 100$  to  $k = 420$ , and isolated spikes appearing at various points. The primary question of interest here is how well the different filters considered eliminate the isolated spikes in this signal while preserving the low-level details, especially the sinusoidal component.

Fig. 2 shows the results obtained for the median filter with  $K = 5$  applied to this signal sequence, which eliminates the noise spikes completely. In the noise-contaminated portion ( $k = 1$  through  $k = 240$ ), this filter strongly suppresses the low-level noise, which may or may not be desirable. In general, the discontinuous transitions are preserved well but the sinusoidal variations are clipped significantly, a highly undesirable distortion in many applications.

Fig. 3 shows the results obtained for the Hampel filter with the same window width ( $K = 5$ ) and the threshold parameter  $t = 1$  applied to the simulated signal sequence shown in Fig. 1. Here, the impulsive noise spikes are removed as effectively as they are by the median filter, but the low-level detail preservation is much better. In particular, note that the low-level noise is largely preserved, and there is no obvious

Figure 2: Response of the median filter with  $K = 5$  to the signal from Fig. 1.

distortion of the sinusoidal component here.

The root sequences for the standard median filter have been well-characterized [3], so it is worth noting that the set  $\mathcal{R}_t$  of root sequences for the Hampel filter with threshold  $t$  contains the median filter root sequence  $\mathcal{R}_0$  for all  $t \geq 0$ . Specifically, if  $s \leq t$ , it follows that:

$$|x_k - m_k| \leq sS_k \leq tS_k \Rightarrow \mathcal{R}_s \subset \mathcal{R}_t. \quad (3)$$

Also, note that there exist non-constant sequences for which  $S_k$  is identically zero. Specifically, if more than  $K$  of the values  $x_{k-j}$  in the moving data window have the same value, this value corresponds to the median  $m_k$  and  $x_{k-j} - m_k = 0$  for these points. As a consequence,  $S_k = 0$ , regardless of the other values in the data window, an effect is known as *implosion* in the statistics literature. Hence, we will refer to sequences  $\{x_k\}$  for which  $S_k$  is identically zero as *implosion sequences*. Their practical significance is that the Hampel filter response is always equal to the corresponding median filter response for these sequences, independent of the threshold parameter  $t$ .

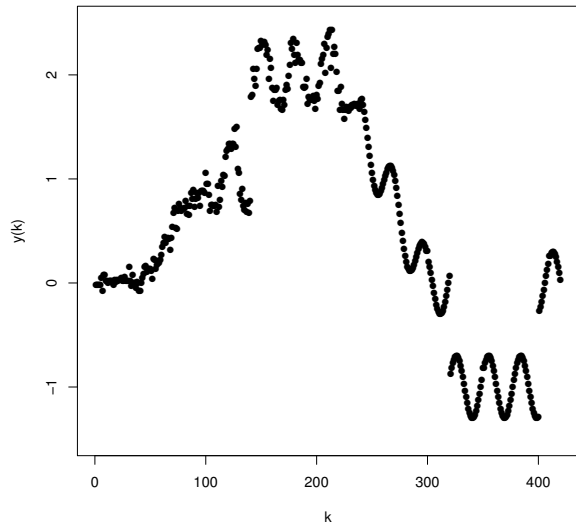


Figure 3: Response of the Hampel filter with  $K = 5$  and  $t = 1$  to the signal from Fig. 1.

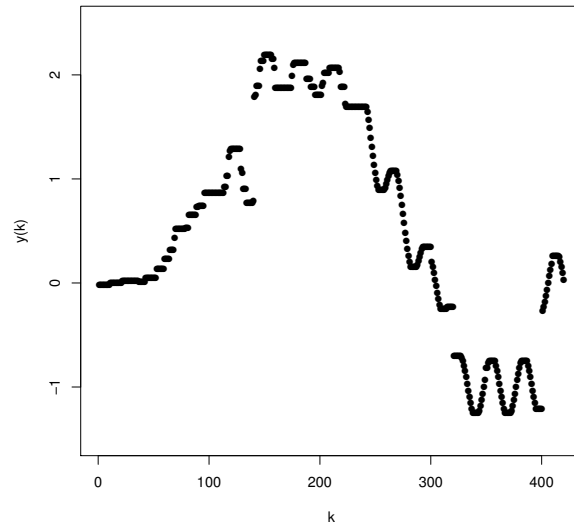


Figure 4: Response of the recursive median filter with  $K = 5$  to the signal from Fig. 1.

### 3 Median filter extensions

The *recursive median filter* replaces the standard median filter's symmetric moving window  $\mathbf{w}_k$  with the recursive data window:

$$\mathbf{r}_k = \{y_{k-K}, y_{k-K+1}, \dots, x_k, \dots, x_{k+K-1}, x_K\}, \quad (4)$$

where  $y_{k-j}$  represents the output at prior time  $k-j$  of the standard median filter applied to the input sequence  $\{x_k\}$ . This extension exhibits a number of interesting properties, including *idempotence* [4]—i.e., a single application of the recursive median filter maps it into the filter's root set. It has also been shown that the root set for the recursive median filter is identical to that for the standard median filter.

*Weighted median filters* replace each element  $x_{k-j}$  in the standard median filter's moving data window  $\mathbf{w}_k$  with its  $w_j$ -fold replicate  $w_j \diamond x_{k-j}$  for some positive integer weight  $w_j$ . This extension increases the median filter's flexibility, but also complicates its analysis: for example, no complete characterization of the root sequences of arbitrarily weighted median

filters is known. The class of *recursive weighted median filters* adopts both of these modifications: the recursive moving window  $\mathbf{r}_k$  defined in Eq. (4), and the  $w_j$ -fold replication of the weighted median filter. Since this filter class includes both of the previous ones as proper subsets, its flexibility is even greater, as is the complexity of its analysis. For a more detailed discussion of both of these filter classes, refer to the survey by Yin *et al.* [7].

Fig. 4 shows the results obtained when the recursive median filter with  $K = 5$  is applied to the simulated data sequence shown in Fig. 1. Comparing this filter response with Fig. 2 for the corresponding nonrecursive filter, it appears that the distortion of the sinusoidal component is more pronounced in the noise-contaminated portion of the signal (i.e., for  $k$  between 100 and 240).

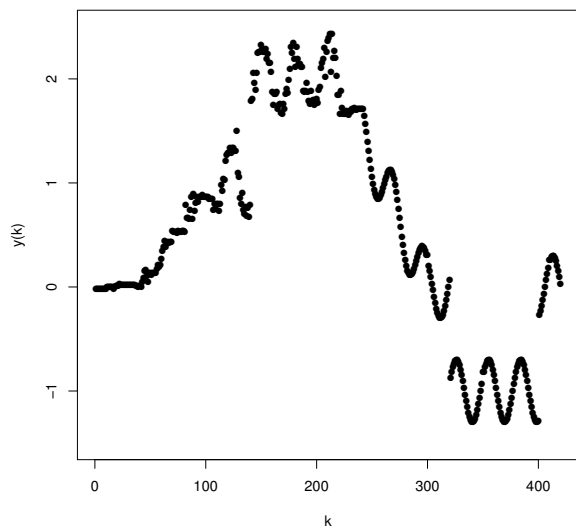


Figure 5: Response of the recursive Hampel filter with  $K = 5$  and  $t = 1$  to the signal from Fig. 1.

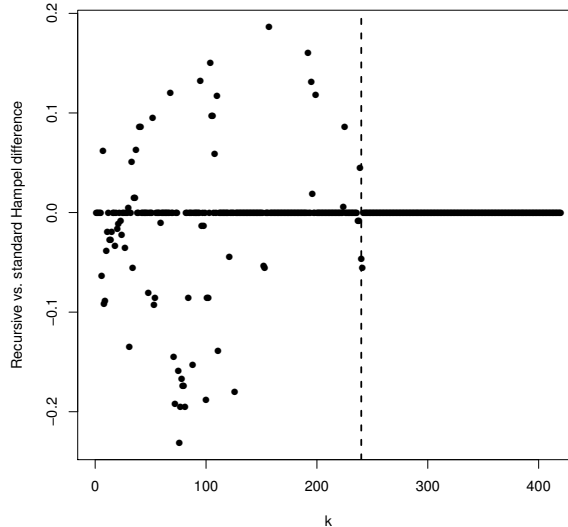


Figure 6: Difference between the recursive and non-recursive Hampel filter responses in Figs. 3 and 5.

## 4 Generalized Hampel filters

The generalized Hampel filters considered here are obtained by replacing the standard median filter used in defining the extensions discussed in Sec. 3 with the Hampel filter described in Sec. 2. Due to space limitations, we only consider the recursive Hampel filter here, comparing it with both its nonrecursive counterpart and the recursive median filter.

Fig. 5 shows the response of the recursive Hampel filter with  $K = 5$  and  $t = 1$  to the signal sequence shown in Fig. 1. Comparing this plot with Fig. 4 for the recursive median filter, it is clear that the recursive Hampel filter preserves the low-level signal details much better than the recursive median filter does, especially for  $k$  between  $\sim 100$  and  $\sim 200$ , where the recursive median filter's distortion is particularly pronounced. Careful comparison of Figs. 5 and 3 suggest that the recursive Hampel filter is more aggressive than the nonrecursive filter, as in the case of the median filter. A more detailed picture of the effects of recursion for both of these filters may be

seen by examining Figs. 6 and 7, which show the differences between the recursive and nonrecursive filter responses for the Hampel filter and the median filter, respectively. In both cases, it is interesting to note that these differences are more pronounced in the presence of noise (i.e., for  $k \leq 240$ ) than in its absence; in particular, there is no difference between the Hampel filter responses in the absence of noise (i.e., to the right of the dashed vertical line). In contrast, for the median filter, the differences are more pronounced in the presence of noise and there are clear differences in the absence of noise.

## 5 Summary

This paper has defined a class of filters obtained by replacing the standard median filter with the Hampel filter in a number of important median filter extensions: the recursive median filter, the weighted median filter, and the recursive weighted median filter. The results presented here are preliminary, based on

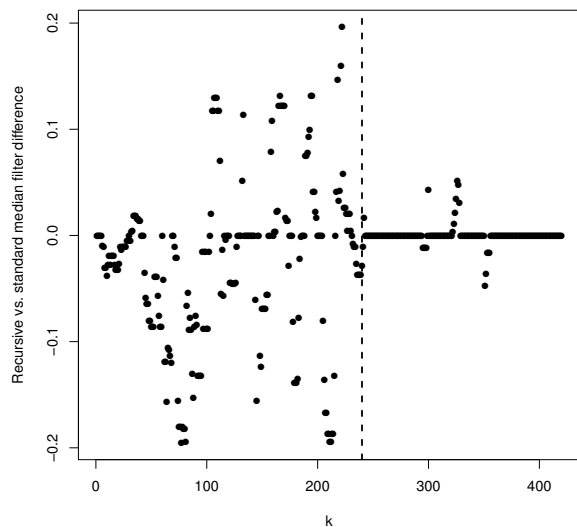


Figure 7: Difference between the recursive and non-recursive median filter responses in Figs. 2 and 4.

limited experimentation with a single example, but they do suggest that the addition of the threshold parameter  $t$  that defines the Hampel filter provides an added degree of design flexibility. An obvious question is how best to exploit this flexibility.

The class of generalized Hampel filters defined here also raises a number of interesting characterization questions. In particular, essentially any characterization result available for median filters and their extensions—e.g., the idempotence of the recursive median filter—leads directly to the more general question: does this result still hold if the median filter on which it is based is replaced with an arbitrary Hampel filter?

In addition, a new class of open questions that is raised by the generalized Hampel filter class described here concerns the implosion sequences introduced in Sec. 2. Since all of the filters in this class are based on the MAD scale, they are all subject to implosion, but the conditions under which this occurs may be very different for the different generalizations of the basic Hampel filter. Thus, it would

be of interest to derive conditions—analogue to the root sequence characterizations for the median filter and its extensions—under which implosion occurs in the different generalized Hampel filter structures.

Finally, it is worth noting that the results presented here are all based on filter implementations in Python, an open-source programming language that is becoming increasingly popular in scientific and engineering applications because it is both free and extremely flexible. Detailed descriptions of median filter implementations in Python, along with all of the median filter extensions described here are given in the forthcoming book by Pearson and Gabbouj [5].

## References

- [1] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*, CRC Press, 1997.
- [2] L. Davies and U. Gather, “The identification of multiple outliers,” *J. American Statistical Assoc.*, vol. 88, 1993, pp. 782–801.
- [3] N.C. Gallagher and G.L. Wise, “A Theoretical Analysis of the Properties of Median Filters,” *IEEE Trans. Acoustics, Speech, Signal Proc.*, vol. 29, 1981, pp. 1136–1141.
- [4] T.A. Nodes and N.C. Gallagher, “Median filters: Some modifications and their properties,” *IEEE Trans. Acoustics, Speech, Signal Proc.*, vol. 30, 1982, pp. 739–746.
- [5] R.K. Pearson and M. Gabbouj, *Nonlinear Digital Filtering with Python*, CRC Press, in press.
- [6] G.L. Sicuranza and A. Carini, “On a class of nonlinear filters,” in *Festschrift in Honor of Jaakko Astola on the Occasion of his 60th Birthday*, I. Tabus, K. Egiazarian, and M. Gabbouj, eds., Tampere International Center for Signal Processing, Tampere, Finland, 2009, pp. 115–143.
- [7] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, “Weighted Median Filters: A Tutorial,” *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 43, 1996, pp. 157–192.