

# KLT Based Quality Controlled Compression of Single Lead ECG

by **T. Blanchett, G. C. Kember, and G. A. Fenton**

Department of Applied Mathematics

Technical University of Nova Scotia

Halifax, Nova Scotia

Canada, B3J 2X4

## Abstract

An ECG compression algorithm based on a combination of the Karhunen-Loeve transform (KLT) and multi-rate sampling is introduced. The use of multi-rate sampling reduces KLT computational times to those reported for wavelet-packet based compression techniques. A beat-by-beat quality controlled compression criterion is shown to be necessary to ensure clinically adequate reconstruction of each beat. The resulting quality controlled algorithm efficiently achieves compression rates of approximately 30-40:1 for the MIT-BIH database.

## 1. Introduction

The electrocardiogram (ECG) is used extensively as a basic diagnostic tool in cardiology. In order to resolve the relatively brief QRS complex, the ECG recording is sampled at frequencies on the order of 1 KHz. ECG compression is desirable because a day's recording typically requires on the order of 100 Megabytes of digital storage.

Recent research [1] employing a small subset of the MIT-BIH arrhythmia database [2] suggests that compression ratios of about 20:1 may be achieved using a wavelet-packet based method. It was argued in [1] that this method allows the maintenance of a clinically acceptable level of error. Furthermore, a naive Karhunen-Loeve transform (KLT) approach was shown, in comparison, to provide a lower average compression rate, and to be computationally slow. This paper describes a KLT based method which offers substantially improved computational efficiency, quality controlled error, and increased compression.

In the KLT approach each of  $N$  beats is represented as an  $M$  dimensional row vector in an  $N \times M$  beat matrix. The KLT basis is taken to be the  $M$  eigenvectors of the beat matrix's autocovariance matrix. Associated with each eigenvector is an eigenvalue which is a measure of the contribution of each eigenvector to the variance of the beat matrix. The KLT is optimal in that it minimizes the mean square error of representation when a subset of transform coefficients associated with the largest eigenvalues are retained.

In particular, the KLT requires fewer transform coefficients to achieve a given error bound than the wavelet approach [1].

In order to reduce the computational burden of the KLT, a multi-rate sampling strategy is adopted, as discussed in Section 2. The error between the original sample and the compression derived reconstruction is investigated in Section 3. This investigation leads to an ‘on-the-fly’ quality control measure for use in error management. One difficulty with standard variance based error measures is that although the average error over a population of beats may be small, the error for an *individual* beat, or portion of a beat, may be large and this can have significant clinical implications. It is shown that, in order to ensure that reconstructed signals can be properly used for clinical evaluation, the error measure must be taken locally.

The resulting method is then applied to the entire MIT-BIH database. It is shown to run at least as fast as that achieved in [1] while providing a two- to four-fold increase in average compression rates and a significantly improved sample to reconstruction comparison.

## 2. Multi-rate Sampling

The ECG sampling frequency is normally chosen to resolve the QRS complex and is set to 360Hz in the MIT-BIH database. However, since the QRS complex operates at comparatively short timescales, the remainder of each beat tends to be oversampled. To assess an appropriate sampling rate within a beat, the R-peak of the QRS complex is first identified. This information is supplied with the MIT-BIH database. Each beat is then divided into three consecutive time windows, or *blocks*, labeled  $\widehat{PQ}$ ,  $\widehat{QRS}$ , and  $\widehat{ST}$ , the hat added to emphasize that these blocks will approximate the standard clinically defined intervals. In the context of the following algorithm, the blocks are used to define time windows containing distinct features of clinical interest. For each beat the R-peak is centered in the  $\widehat{QRS}$  block and the block’s duration is set to 1/4 second (where the typical normal QRS complex is about 50-110 msec [3]). The PR interval is generally less than 40 percent of its corresponding RR interval [3]. The duration of the  $\widehat{PQ}$  block is therefore set to 40 percent of the RR-interval less one-half of the duration of the  $\widehat{QRS}$  block. Similarly, the duration of the  $\widehat{ST}$  block is set to the 60 percent of its corresponding RR-interval less one-half the  $\widehat{QRS}$  block. The resulting blocks,  $\widehat{PQ}$ ,  $\widehat{QRS}$ , and  $\widehat{ST}$ , can be approximately associated with the three major phases of the cardiac cycle, that is, atrial activation, ventricular activation and ventricular repolarization respectively [3].

An appropriate sampling rate for each block may be derived from its average spectral density function (SDF). The average SDF is constructed by randomly sampling 200 beats from the MIT-BIH population, estimating the SDF within each block of each beat, and averaging across the ensemble. An example of the

resulting average SDF's computed for normal beats is shown in Figure 1. Note that the  $\widehat{\text{PQ}}$  and  $\widehat{\text{ST}}$  blocks have negligible power over about 10 Hz, while the  $\widehat{\text{QRS}}$  block operates at up to about 30 Hz. The  $\widehat{\text{QRS}}$  block can thus be conservatively downsampled at a rate of 180 Hz (from the original 360 Hz), providing a maximum frequency resolution of 90 Hz. An extensive investigation [7] achieved over 91% detection and identification rates of premature ventricular contractions using a sampling rate of 90 Hz. The average SDF's for the  $\widehat{\text{PQ}}$  and  $\widehat{\text{ST}}$  blocks indicate that these regions can be conservatively downsampled at a rate of 72 Hz (every 5'th point), providing a frequency resolution of up to 36 Hz. These sampling rates are more than adequate to preserve details of clinical interest in the original signal.

A well known side effect of downsampling is aliasing [4]. To minimize this phenomenon, a computationally simple 9-point centered local average is used at each downsampling point in the  $\widehat{\text{PQ}}$  and  $\widehat{\text{ST}}$  blocks and a 3 point centered local average at each downsampling point in the  $\widehat{\text{QRS}}$  block. The centered average acts as a zero-phase, low-pass filter which attenuates noise (less than 10% passing) above 30 Hz in the  $\widehat{\text{PQ}}$  and  $\widehat{\text{ST}}$  blocks and above 90 Hz in the  $\widehat{\text{QRS}}$  block. Since both of these values approximate the Nyquist frequencies corresponding to the downsampling rates there will be negligible aliasing taking place nor is the signal of interest compromised. Because the block boundaries are chosen to be far removed from the physiologically rapid QRS interval, discontinuities at block boundaries will be negligible.

### 3. Quality Controlled Compression

Clinical use of the ECG recording is ultimately based on a beat-by-beat analysis of features of the PQ, QRS and ST portions within each beat. However, the accuracy of transform-based compression algorithms in the literature has been evaluated using ensemble averages of the mean square error of beat reconstruction ([1],[5],[6],[7]). Ensemble averages obviously do not ensure clinically adequate, beat-by-beat reconstruction since a number of beats must exceed the error criterion. Most importantly, ensemble averages do not ensure capturing clinically relevant features in the PQ, QRS, and ST portions, since each portion makes widely differing contributions to the variance.

Hence, the aim of quality controlled compression is to ensure clinically adequate reconstruction on a beat-by-beat basis, preserving relevant features within each beat. To implement quality controlled compression, a criterion is proposed herein in which the number of retained transform coefficients is such that *the relative mean square errors,  $\epsilon_j$ , for each block  $j = 1, 2, \text{ or } 3$  (for blocks  $\widehat{\text{PQ}}$ ,  $\widehat{\text{QRS}}$ , and  $\widehat{\text{ST}}$  respectively) are less than or equal to a prespecified tolerance.* Specifically, if the relative error over block  $j$  is defined as

$$\epsilon_j = \frac{\|\hat{\mathbf{X}}_j - \mathbf{X}_j\|_2^2}{\|\mathbf{X}_j\|_2^2}, \quad j = 1, 2, 3$$

where  $\|\cdot\|_2$  is the 2-norm,  $\mathbf{X}_j$  is the original sample vector over block  $j$ , and  $\hat{\mathbf{X}}_j$  is its reconstruction from the retained transform coefficients, then the number of retained coefficients for the entire beat is determined to be the minimum integer such that  $\max_j \epsilon_j \leq b$  for some tolerance  $b$ . This criterion is iterative since  $\epsilon_j$  depends on the number of retained coefficients. However, since  $\hat{\mathbf{X}}_j$  is easily constructed incrementally, the minimum number of retained coefficients is found efficiently. Note, that it is possible to consider placing each block as a row of the ‘beat’ matrix, leading to a KLT basis for *each block*. This approach was not investigated since it would still require the quality control criterion and it was deemed to be less computationally efficient without a corresponding significant increase in accuracy and/or compression ratio.

#### 4. Implementation Details

The 48 records of the MIT-BIH database are analyzed unabridged in ten minute segments. Baseline drift in each segment is removed by subtracting a 1-second wide centered moving average. Each segment is then partitioned into  $N$  beats and each beat is loaded sequentially as a row vector of dimension  $M$  into a beat matrix. The blocking of a segment is defined as follows: (1) the maximum RR-interval  $R_{max}$  is identified, (2) the number of data points in the  $\widehat{PQ}$ ,  $\widehat{QRS}$ , and  $\widehat{ST}$  blocks,  $N_{PQ}$ ,  $N_{QRS}$ , and  $N_{ST}$  respectively are; (a)  $N_{PQ} = (0.4R_{max} - 1/8)/\Delta_{PQ}$ , (b)  $N_{QRS} = (1/4)/\Delta_{QRS}$ , (c)  $N_{ST} = (0.6R_{max} - 1/8)/\Delta_{ST}$ , where  $\Delta_{PQ}$ ,  $\Delta_{QRS}$  and  $\Delta_{ST}$  are the sample time increments in the  $\widehat{PQ}$ ,  $\widehat{QRS}$  and  $\widehat{ST}$  blocks respectively (after downsampling). The row dimension of the beat matrix is set to  $M = N_{PQ} + N_{QRS} + N_{ST}$ , which is determined by  $R_{max}$ . Since all remaining beats have RR-intervals less than  $R_{max}$ , they must be aligned within the matrix. To do this, the R-peak for each beat vector is placed at the index  $N_{PQ} + (N_{QRS} - 1)/2$  (for  $N_{QRS}$  odd) and padding is provided before the samples (using the first sample in the  $\widehat{PQ}$  block) and after the samples (using the last sample in the  $\widehat{ST}$  block) to complete the vector.

#### 5. Results

The following results are based on the channel 1 lead of the MIT-BIH database where it is assumed that the number of bits per value on input is unchanged on output. To obtain meaningful compression measures, the total storage required by the retained projection coefficients must be considered. In the following it is assumed that the precision at which the required coefficients are stored is the same as that for the original signal. If the dimension of the beat matrix is  $N > M$  (true if no beat lasts more than about 5 seconds throughout the 10-minute segment), then the KLT produces an  $M \times M$  matrix of eigenvectors, an  $N \times M$  matrix of transform coefficients, and the  $M$  eigenvalues  $\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2$ .

For the variance criterion,  $m$  eigenvalues (arranged in descending order of magnitude) and eigenvectors, along with  $m$  transform coefficients for each of  $N$  beats are retained with  $m$  chosen so that  $\sum_{i=1}^m \sigma_i^2 / \sum_{i=1}^M \sigma_i^2 \geq \delta_m$ . Including the  $N$  RR-interval values needed to reconstruct each beat, the variance control criterion requires the storage of  $m(M + 1) + N(m + 1)$  values in total. A 10 minute recording originally containing  $(10)(60)(360) = 216,000$  data values typically has  $N = 750$  (corresponding to a heart rate of 75 beats per minute) and  $M = 100$  after downsampling. For  $m = 3$  retained transform coefficients, the storage requirement totals  $3(100 + 1) + 750(3 + 1) = 3303$  values yielding a compression ratio  $C = 65:1$ . If  $m = 10$ ,  $C$  drops to about 23:1.

In contrast to the variance criterion, the quality controlled criterion, discussed in Section 3, requires the storage of; 1) a variable number of transform coefficients  $m_i, i = 1, 2, \dots, N$ , 2)  $m_{MAX} = \max_i m_i$  eigenvalues and eigenvectors, and 3) the RR-interval for each of  $N$  beats. The total storage requirement for segment of  $N$  beats is then  $m_{MAX}(M + 1) + N(\bar{m} + 1)$  where  $\bar{m}$  is the average number of retained transform coefficients over the  $N$  beats. The compression ratio for a 10 minute segment (as above) with  $N = 750$ ,  $M = 100$ ,  $\bar{m} = 3$  and  $m_{MAX} = 20$  is then  $C = 43:1$ . While most beats are stored using only 2 or 3 transform coefficients this approach allows for the occasional (pathological) beat to be accurately represented using a larger number of transform coefficients, maintaining, overall, an excellent compression ratio.

The average compression ratio over the MIT-BIH database for the variance criterion with  $\delta_m = 0.98$  is  $C \approx 53$  while the same for the quality control criterion with tolerance  $b = 0.15$  is  $C \approx 32$ . These numbers for the variance and quality control criteria are base compression ratios; they can be improved 35% and 10%, respectively, by using only 6 bits to store transform coefficients, as in [1]. In Figure 2, the compression achieved for each ten minute segment over the 48 records of the MIT-BIH database is reported. Typically, the quality control criterion provides a lower average compression ratio over each ten minute segment. This implies, however, that there are a significant number of beats that meet the variance criterion but go on to fail the quality control criterion (so that  $m$  must be increased).

Figure 3 examines the effect of failing the quality control criterion over a segment where the variance criterion required  $m = 3$  for  $\delta_m = 0.98$ . Some reconstructed beats that required more transform coefficients for quality control are shown alongside the reconstruction taken from the  $m = 3$  coefficients specified by the variance criterion. Three of the beats are represented adequately for clinical diagnosis by both methods. However, the variance criterion reconstruction completely misses the depressed ST segment of the second beat. From a clinical point of view, this sporadic loss of potentially significant features is clearly not acceptable. In contrast, the quality control criterion, while requiring 14 transform coefficients for the second beat, preserves the behavior of the signal.

In summary, an algorithm employing KLT, along with its attractive properties relating to error measures and best bases, combined with downsampling and a quality control criterion has been found to lead to efficient and accurate compression of ECG recordings. By careful consideration of the bandwidth containing the signal of interest, downsampling can be performed without loss of pertinent information. Downsampling, in turn, leads to efficient computation of the KLT, significant eigenvector storage savings, and compression ratios which typically exceed those achieved via wavelet packet-based algorithms. The KLT also naturally lends itself to noise model extensions relevant to signal identification. Finally, the quality control criterion allows clinically acceptable reconstructions while maintaining high compression ratios.

## 6. Acknowledgements

Thanks for financial support are due to the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grants OGP0155608 and OGP0105445 and under the NSERC Postgraduate Scholarship Program. Any opinions, findings, and conclusions and recommendations are those of the authors and do not necessarily reflect the views of the aforementioned organizations.

## 7. References

- [1] B. BRADIE, "Wavelet packet-based compression of single lead ECG," *IEEE Trans. Biomed. Eng.*, vol. 43, pp. 493–501, 1996.
- [2] G.B. MOODY, and R.G. MARK, "The MIT-BIH arrhythmia database on cd-rom and software for use with it," *Computers in Cardiology 1990*, IEEE Computer Society Press, pp. 185–188, 1991.
- [3] M.J. GOLDMAN, *Principles of Clinical Electrocardiography 8th ed*, Lange Medical Publications, Los Altos, CA, 1973.
- [4] P.P VAIDYANATHAN, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," *Proceedings of the IEEE*, vol. 78, pp. 56–93, 1990.
- [5] A. ENIS CETIN, HAYRETTIN KOYMEN, and M CENGİZ AYDIN, "Multichannel ECG data compression by multirate signal processing and transform domain coding techniques," *IEEE Transactions on Biomedical Engineering*, vol 40, pp. 495–499, 1993.
- [6] G. STRANG, and T. NGUYEN, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley Ma, 1996.

- [7] N. AHMED, and K.R. RAO, *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag, Berlin, 1975.
- [8] G.B. MOODY, and R.G. MARK, “QRS morphology representation and noise estimation using the Karhunen-Loeve transform,” *Computers in Cardiology 1989* , IEEE Computer Society Press, Los Alamitos, CA, pp. 269–272, 1989.

## List of Figures

- Figure 1.** Average block specific spectral density functions.
- Figure 2.** Compression ratios over MIT-BIH database for variance and quality control criteria.
- Figure 3.** Selected beats from Record 100, Segment 3 along with variance and quality control criteria reconstructions.

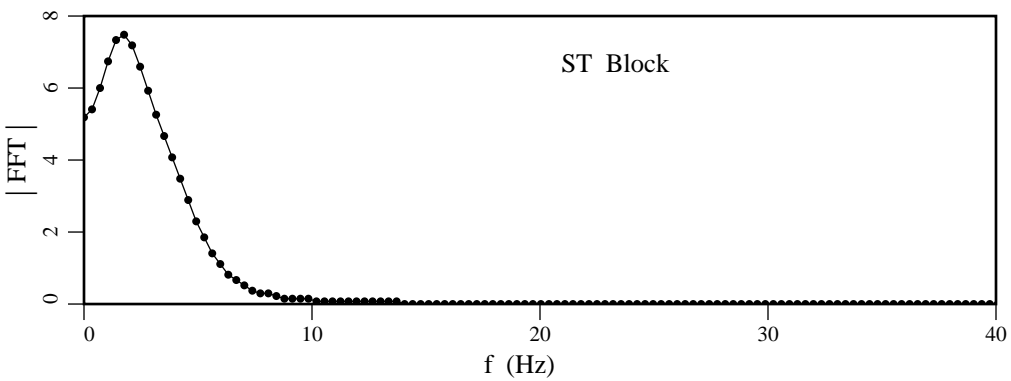
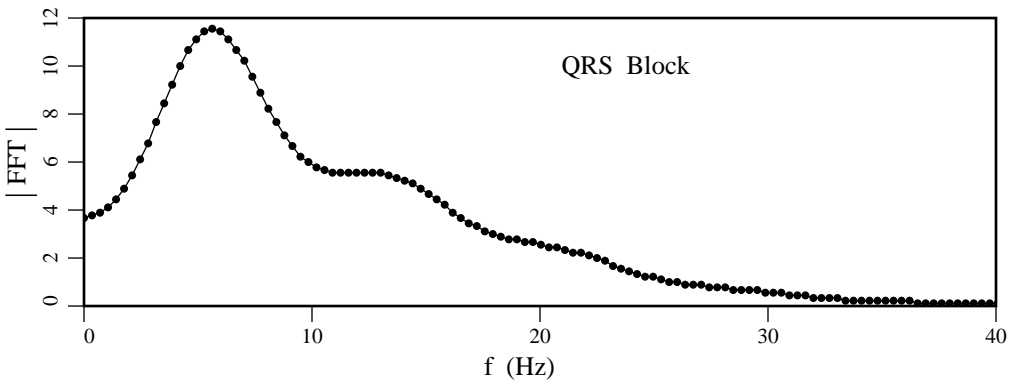
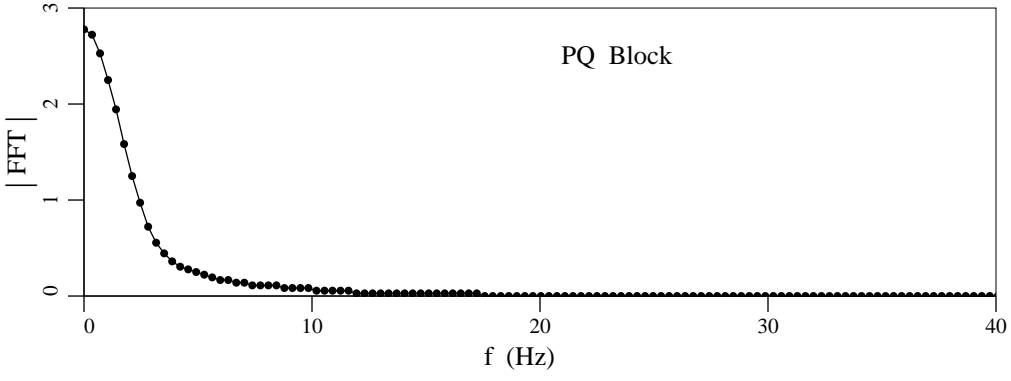


Figure 1 (Blanchett/Kember/Fenton)

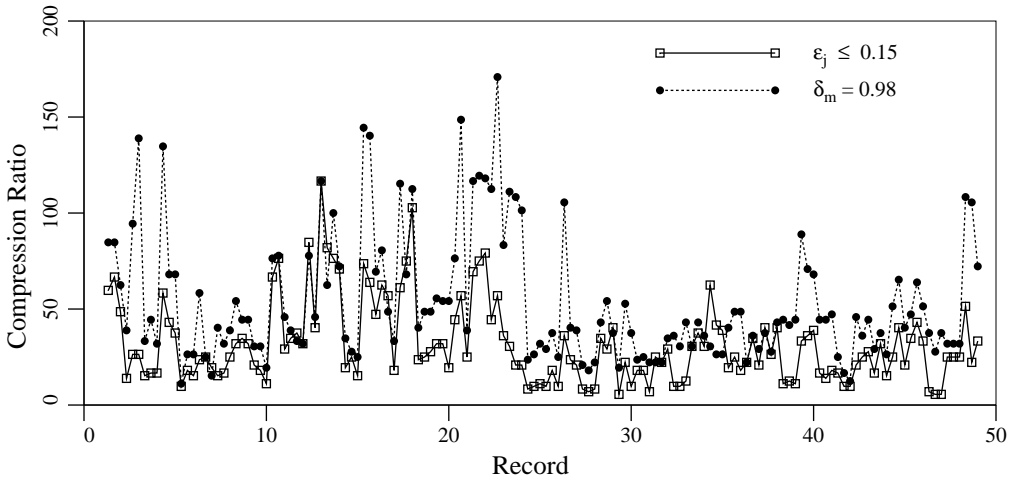


Figure 2 (Blanchett/Kember/Fenton)

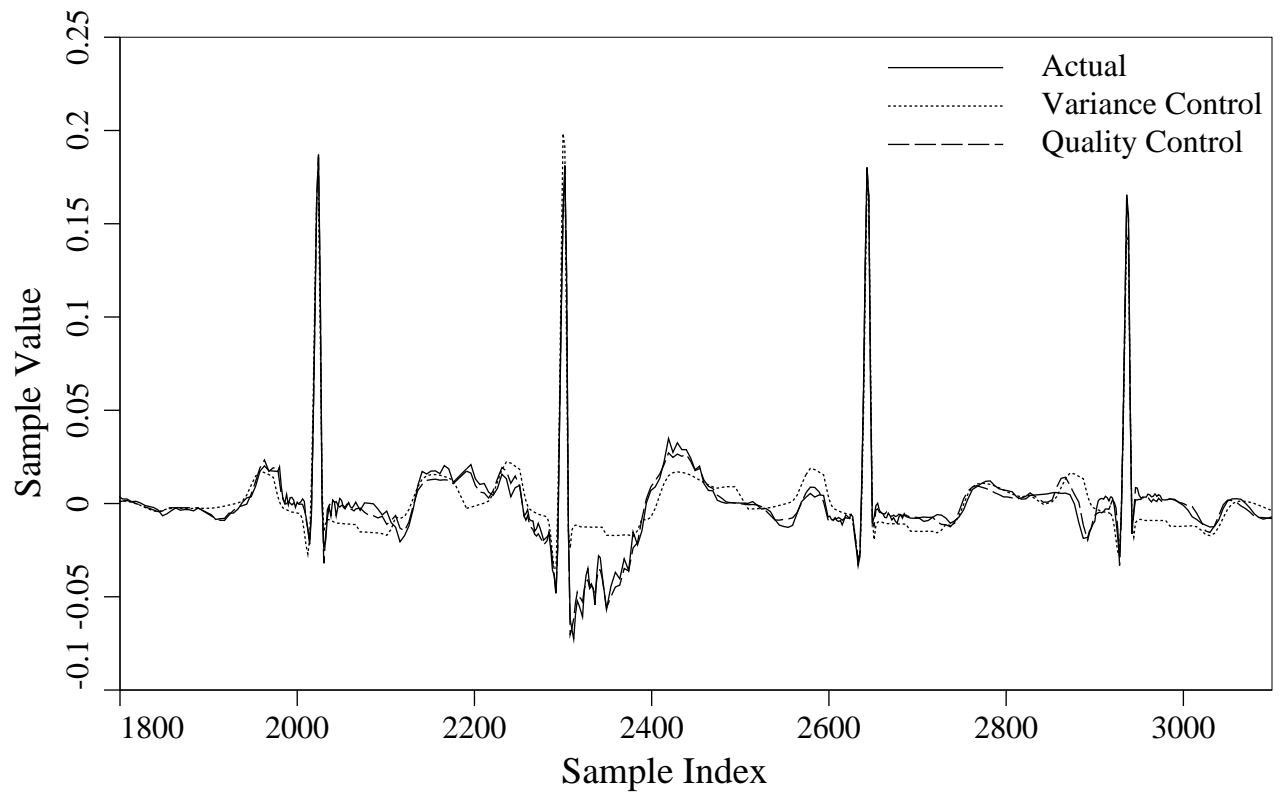


Figure 3 (Blanchett/Kember/Fenton)