Mathematical Techniques for Mitigating Alarm Fatigue

Alarm Fatigue

> "Hospital staff are exposed to an average of 350 alarms per bed per day, based on a sample from an intensive care unit at the Johns Hopkins Hospital in Baltimore."[1]

From the same survey, almost 9 in 10 hospitals indicated they would increase their use of patient monitoring, particularly of Capnography and pulse oximetry, if false alarms could be reduced. [2]

> "Of those hospitals surveyed that monitor some or all patients with pulse oximetry or Capnography, more than 65 percent have experienced positive results in terms of either a reduction in overall adverse events or in reduction of costs."[3]

The problem with attenuating alarm data is achieving the balance between communicating the essential, patient-safety specific information that will provide proper notification to clinical staff while minimizing the excess, spurious and non-emergent events that are not indicative of a threat to patient safety. In the absence of contextual information, the option is usually to err on the side of excess because the risk of missing an emergent alarm or notification carries with it the potential for high cost (e.g.: patient harm or death).

The purpose of this study is to look at the mathematics and some of the techniques and options available for evaluating real-time data. The objective is to suggest a dialog for further research and investigation into the use of such techniques as appropriate. Clearly, patient safety, regulatory, staff fatigue and other factors must be taken into account in terms of aligning on a best approach or practice (if one can even be identified). These aspects of alarm fatigue are intentionally omitted from the discussion at this point (to be taken up at another time) so that a pure study of the physics of the parameter data and techniques for analyzing can be explored.

Patient Controlled Analgesia

Based on reports made to the FDA between 2005 and 2009 [4]:

> "…more than 56,000 adverse events and 700 patient deaths were linked to patient-controlled analgesia (PCA) pumps. One out of 376 post-surgical patients are harmed or die from errors related to… [PCA] … that help relieve pain after surgical procedures."

Slightly more than 70% of hospitals surveyed indicated they would prefer a "single indicator that accurately incorporates key vital signs, such as pulse rate, SpO_2, respiratory rate and $etCO_2$" [5].

But, 19 in 20 hospitals indicate they are concerned with alarm fatigue and almost 9 in 10 hospitals indicated that a reduction in false alarms would likely increase the use of patient monitoring devices such as the pulse oximeter or capnograph.

Alarm reporting from medical devices is key to alerting clinical staff of events. Yet, the concomitant fatigue of responding to many false alarms may render clinical staff "snow-blind" to real events, or cause the alarms to be ignored or even turned off, obviating any potential benefit.

Modeling Discrete Data

To illustrate the scope of the issues faced, a hypothetical sampling of data based on experiential measurements, will be used as the target for discussion. Figure 1 plots simulated end-tidal CO2 versus time for a hypothetical patient. The range and behavior of the data are based on real-data, inclusive of the aberrations (e.g.: spikiness) and trends of such data. The range of values are also span normal and abnormal as well as emergent ranges. In general, normal values for etCO2 span the range of 35-45 mmHg in adult humans. The data shown in this figure fall considerably outside of this range as many capnograph devices have alarm settings for emergent conditions set to provide notifications below 25 mmHg. Depending on clinical workflow and organizational policies, sometimes ranges of conditions are identified whereby yellow alerts (amber alerts) are defined for conditions in which end-tidal CO2 drops between 25 and 15 mmHg, and red alerts for end tidal CO2 below 15 mmHg, or severe hypocapnia.

On the upper-bound side, ranges of end-tidal CO2 above 55 mmHg are sometimes defined by policy as emergent levels of severe hypercapnia.

The cautionary and emergent ranges for the data set under consideration are illustrated in Figure 2.

When considering continuous monitoring of end-tidal CO2, if alarm ranges on the capnograph are as identified in the figure, one can see the potential to issue alarms quite frequently. However, many of these alarms are "one-and-done." That is, they occur because of some aberrant behavior (e.g.: shifting of nasal cannula, or patient moving in bed) that cause the measurement to spike or register as an out-of-bounds value with respect to the alarm levels set on the monitor itself.

If these alarms are issued at the time they occur, lacking context, one may see that they can provide a frequent source of distraction, particularly for the nursing staff.

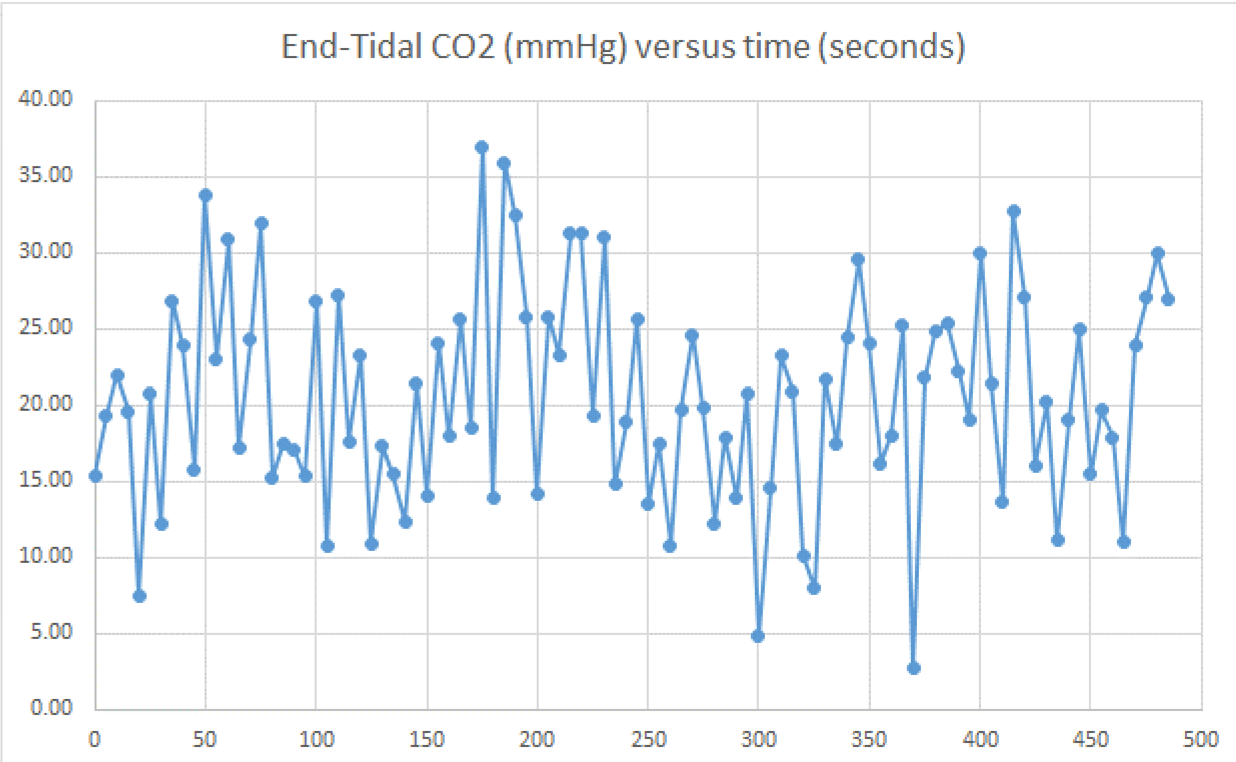

FIGURE 1: MODEL OF END-TIDAL CO2 VERSUS TIME.

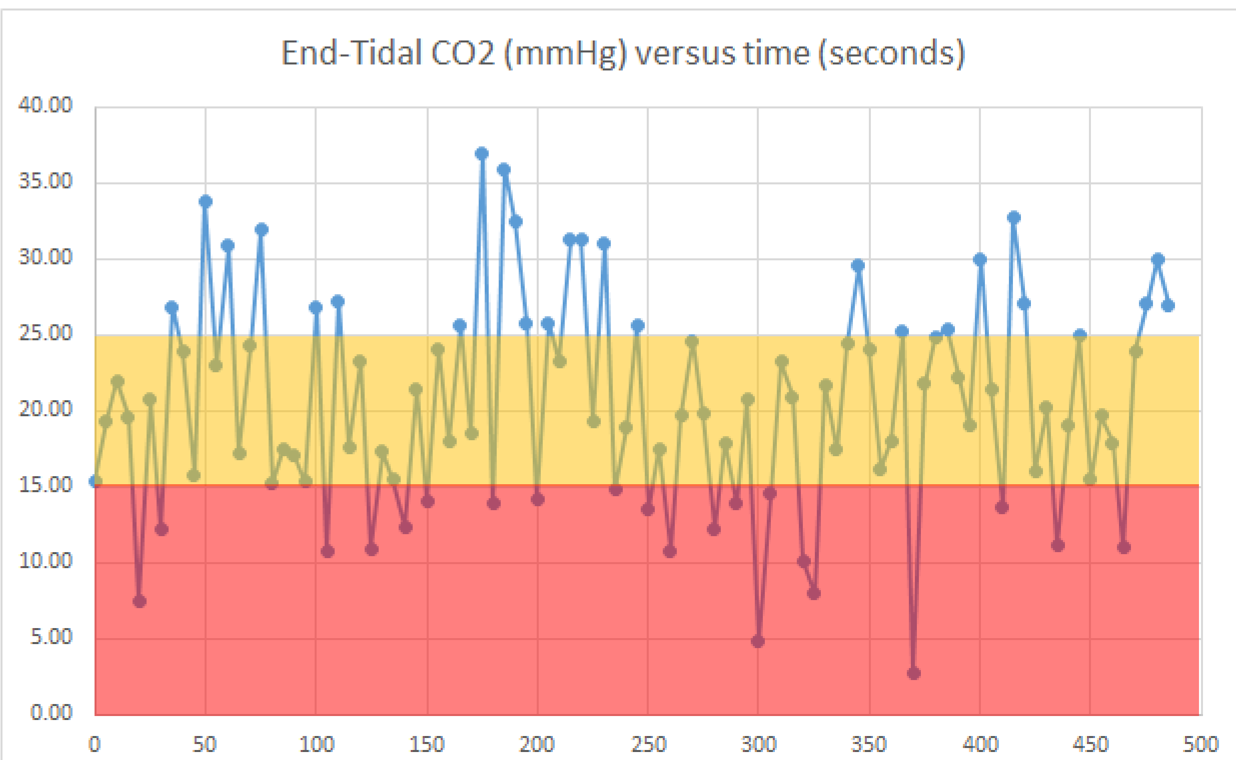

FIGURE 2: CAUTIONARY AND EMERGENT RANGES OF HYPOCAPNIA DISPLAYED ON TOP OF ETCO2 PLOT.

This is not to suggest that the chief objection is irritation to the clinical staff. But, from a patient safety perspective, when alarms go off and there is no clear distinction between an emergent alarm versus a nuisance alarm, this can impact the ability to react appropriately to the patient.

Consider the figures above. Upon inspection it can be seen quite readily that a number of "spikes" wherein measurements are shown dipping down into the "red" area. It can be noticed, too, that in many instances the measurements rise back into a less emergent or normal area upon the very next measurement. Hence, the measurements that would normally trigger the alert do not persist. What is the cause of this? Many possible answers, from artifact in the measurement cannula to movement of the patient.

Problems may be more readily identified if other corroborating information can be brought to bear, such as corresponding changes in respiratory rate, heart rate, SpO2 values. Moreover, if problems are truly present, it would be logical to conclude, based on experience and policy that these "events" would either persist, or would increase in frequency over time. In other words, to verify that behavior are not merely incidental but are correlated to some behavior, the expectation of continued depression or trending depression towards hypocapnia or hypercapnia would be present. A simple way of measuring this (given no other information) is a sequence of measurements at or around the emergent value. For instance, multiple measurements over a period of time of, say, 20 or 30 or 40 seconds in which the values are depressed or elevated. Or, a series of spikes that occur rapidly over a fixed period of time. In reviewing the data retrospectively, and in light of the desire to reduce spurious notifications, several methods will be considered for determining the viability of reducing or filtering out such measurements.

Time Averaging

Each measurement is taken every 5 seconds from the capnograph. Hence, 2 measurements cover 10 seconds' time. A total of 3 measurements cover 20 seconds, and 4 measurements cover 30 seconds' time. An overlay of the time-averaged signals is shown in Figures 3, 4, and 5, respectively.

$$A(t_k) = \sum_{i=k-N}^{i=k} f(i) \times A(t_i)$$

Where $f(i)$ is the fractional weight associated with each measurement $A(t_i)$, where the measurements in the past run from i=k-N, N arbitrary, to i=k. In the specific case of 10 second signal averaging, N=3 (t = 0, 5, 10); in the case of 20 second signal averaging, N=5 (t=0, 5, 10, 15, 20); in the case of 30 seconds signal averaging, N=7 (t=0, 5, 10, 15, 20, 25, 30).

For equal weighting of the measurements, $f(i) = \frac{1}{N}$, and $\sum_{i=k-N}^{i=k} f(i) = 1$.

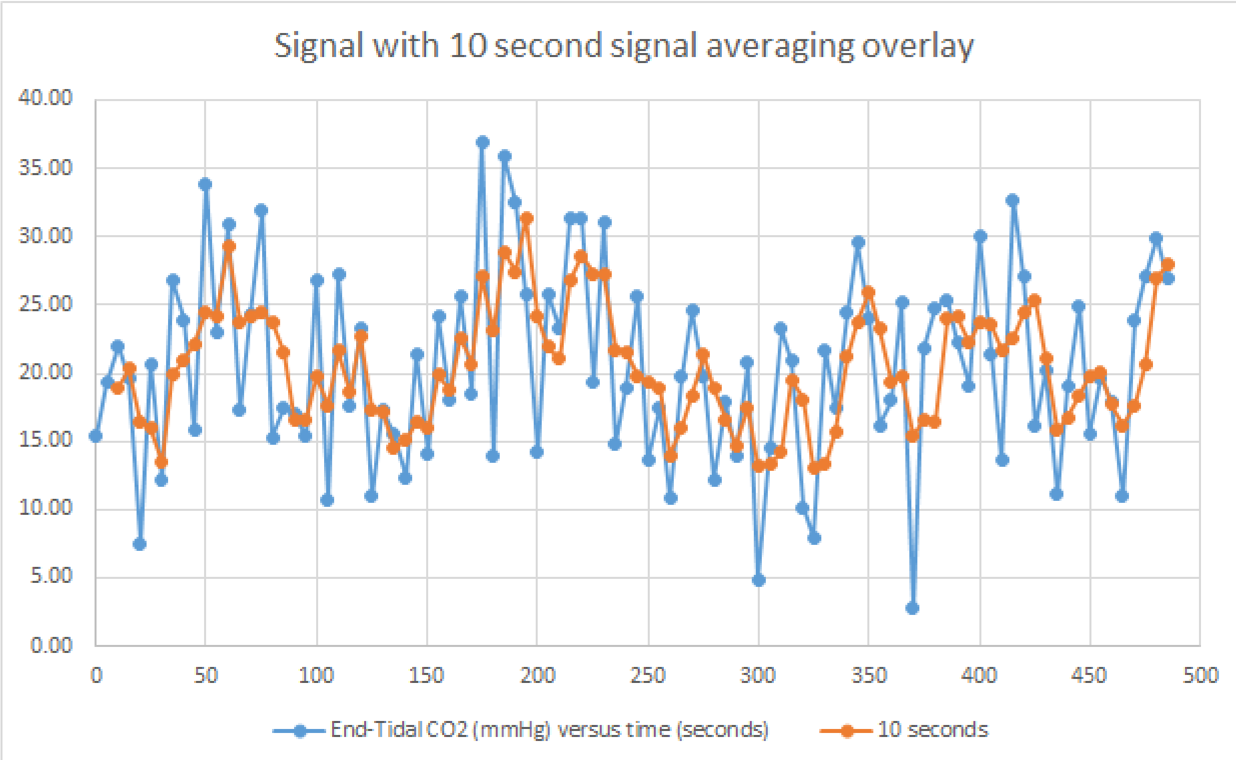

FIGURE 3: SIGNAL WITH 10 SECOND TIME AVERAGING OVERLAY.

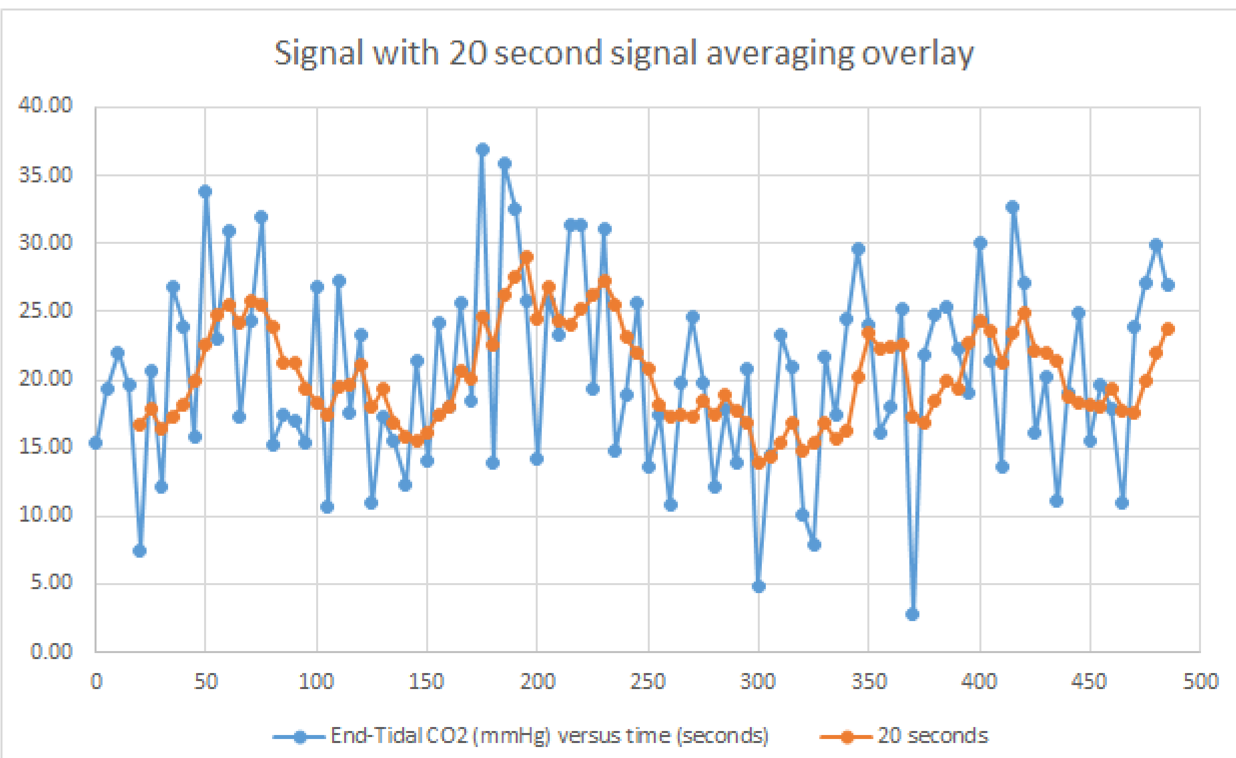

FIGURE 4: SIGNAL WITH 20 SECOND TIME AVERAGING OVERLAY.

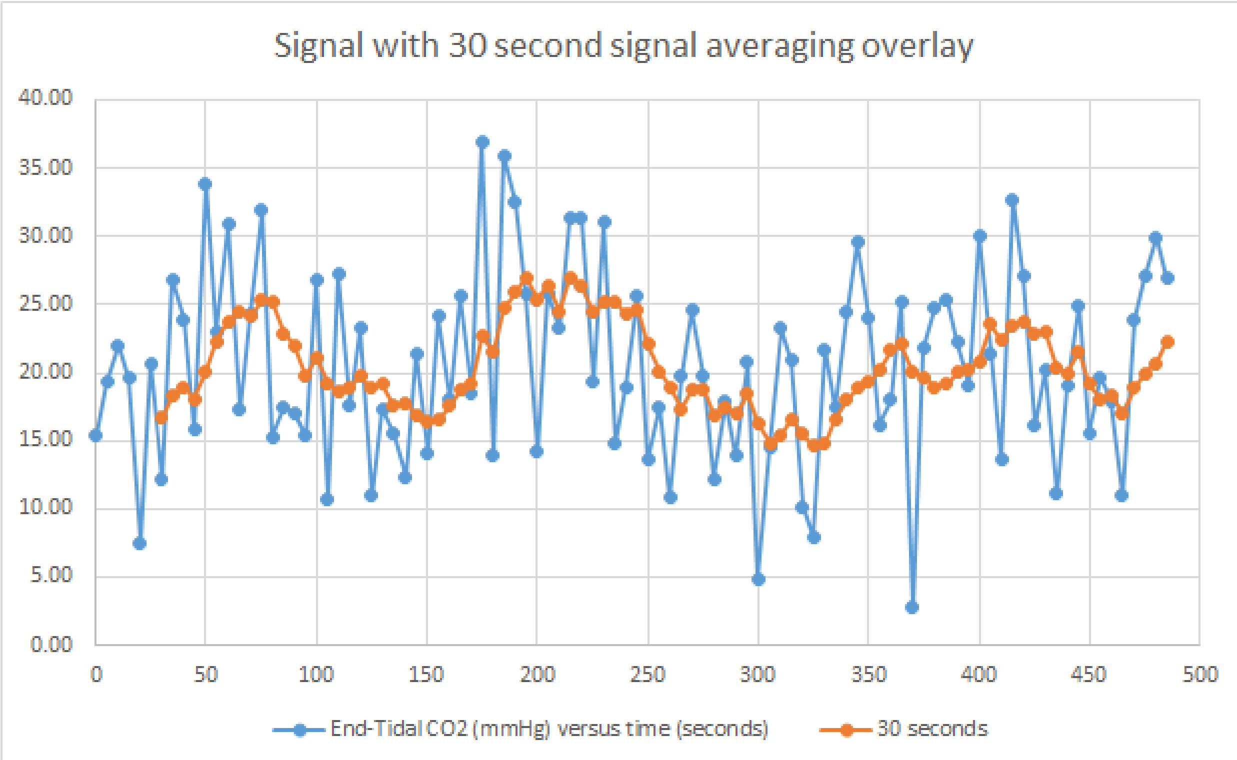

FIGURE 5: SIGNAL WITH 30 SECOND TIME AVERAGING OVERLAY.

Alternative Measurement Weighting

The fractional weight need not be taken as equal across all measurements: the weights can be skewed so that certain measurements have more influence on the overall average. For instance, in the plot of Figure 6, a 10 second averaging with weights as follows is shown:

$$f_{k-2} = 0.15,$$

$$f_{k-1} = 0.25,$$

$$f_k = 0.60$$

This places the most emphasis on the last measurement, deemphasizing the next-to-last measurement and placing even less weight on the first measurement of the series. As a result of this, the value of the time-average is closer to the value of the last raw data measurement, implying that the conditions of the last measurement are to be emphasized over any other.

In contrast, Figure 7 shows the result of transposing the next-to-last and last measurement, placing emphasis on the middle measurement. The weighting for this case is:

$$f_{k-2} = 0.15,$$

$$f_{k-1} = 0.60,$$

$$f_k = 0.25$$

The middle measurement is given the most emphasis in this case.

Regardless of weighting scheme, the selection of an appropriate weighting for individual measurements is left to those to define as partially the subject of academic investigation and partially the empirical assessment of researchers seeking to learn whether an optimal outcome can be defined or preferred. The variation is illustrated here simply to expose the reader to the possibility.

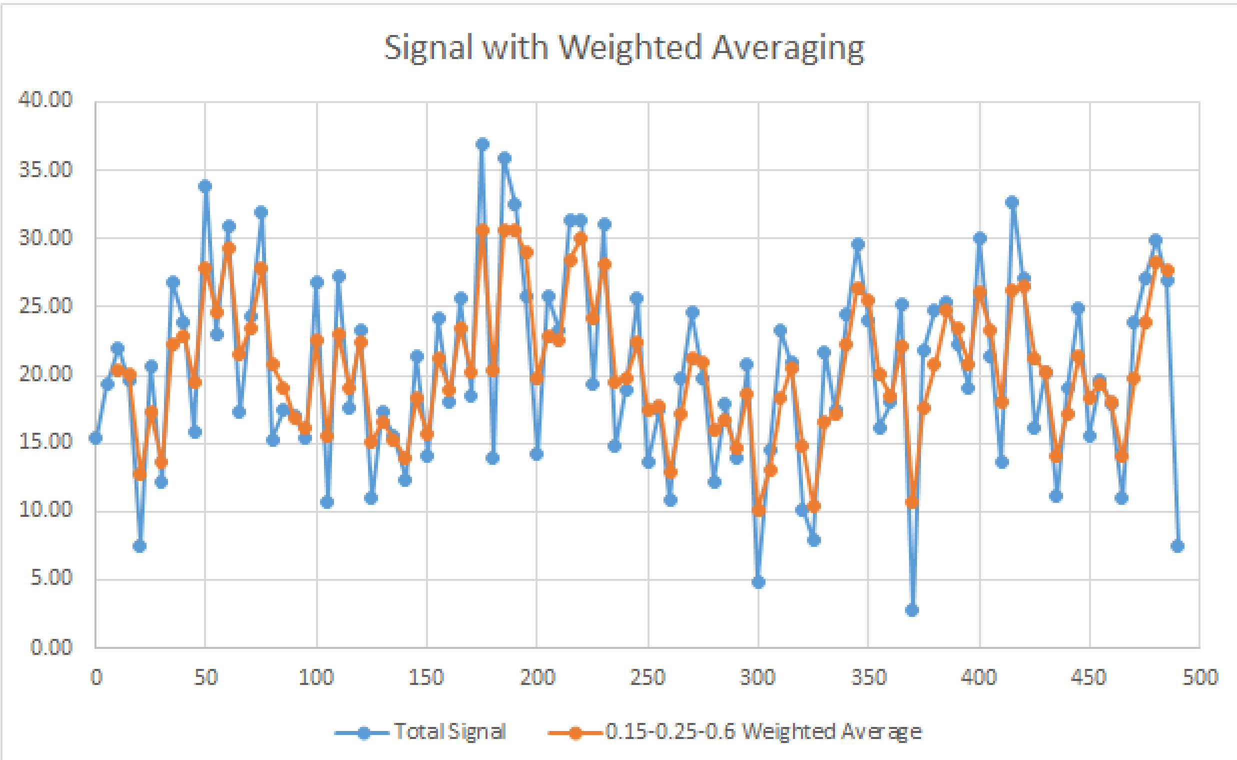

Figure 6: 10 second time averaging with measurement weighting 0.15-0.25-0.60

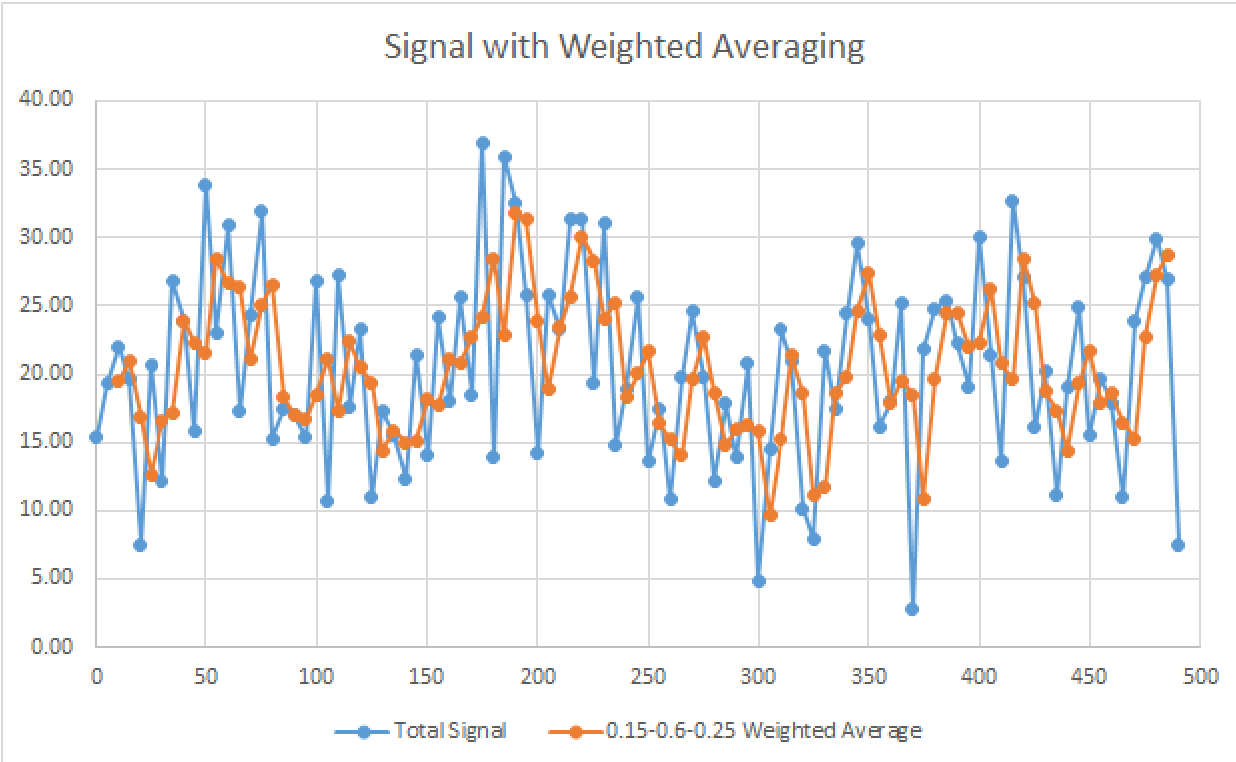

FIGURE 7: 10 SECOND TIME AVERAGING WITH MEASUREMENT WEIGHTING 0.15-0.60-0.25

More Formalized Time-Series Filtering: The Extended Kalman Filter

The concept of optimal filtering of data has many advocates and many applications. The use of formalized methods, including least-squares techniques and the application of Kalman filtering to the study of medical signals and other data has been widely published [6][7][8].

Kalman filtering employs a recursive algorithm that is an optimal estimator to infer and establish an estimate of a particular parameter based upon uncertain or inaccurate observations. A benefit of the Kalman filter is that it allows recursive processing of measurements in real-time as observations are made and, so, can be applied to live data readily. The Kalman filter also provides for tuning and filtering to enable removal or attenuation of noise.

The generalized equations defining the Kalman filter are the state estimate and the measurement update:

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$

- x_k is the state vector containing the estimate of the future state at time k based on previous state at time k-1;
- u_t is the vector containing any control inputs;

- A is the state transition matrix which applies the effect of each system state parameter at time k-1 on the system state at time k;
- B is the control input matrix which applies the effect of each control input parameter in the vector u_k on the state vector;
- w_{k-1} is the vector containing the process noise terms at time k-1 for each parameter in the state vector. The process noise is assumed to be drawn from a zero mean multi-variate normal distribution with covariance given by the covariance matrix Q_k;
- z_k is the vector of measurements at time k;
- H is the transformation matrix that maps the state vector parameters into the measurement domain; and,
- v_k is the vector containing the measurement noise terms for each observation in the measurement vector. The measurement noise is assumed to be zero mean Gaussian noise with covariance R.

The filter solution balances the confidence in the measurements with the confidence in the estimates. That is, the filter will respond more closely to the measurements if the "belief" or confidence in the measurements is greater than the confidence in the estimates, and vice versa. If the measurement noise is high, the confidence in the measurements is fairly low, and the filter will smooth out the transitions between measurements, resulting in a state estimate which is less perturbed but also that does not react to sudden changes in measurements (hence, less likely to react to sudden or spurious changes).

The solution process is as follows:

Time Update (Prediction)	Measurement Update (Correction)
$\hat{x}_{k\|-} = A\hat{x}_{k-1} + Bu_k$ $P_{k\|-} = AP_{k-1}A^T + Q$	$K_k = P_{k\|-}H^T\left(HP_{k\|-}H^T + R\right)^{-1}$ $\hat{x}_k = \hat{x}_{k\|-} + K_k\left(z_k - H\hat{x}_{k\|-}\right)$ $P_k = (I - K_kH)P_{k\|-}$

The solution proceeds with an initial guess on the covariance (P) and the state estimate $\left(\hat{x}_k\right)$.

The minus sign (k|-) indicates the estimate at the previous iteration before update occurs. When the assumptions and particulars of the signal are applied to these specific equations, they reduce to the following:

Time Update (Prediction)	Measurement Update (Correction)
$\hat{x}_{k\|-} = \hat{x}_{k-1} + u_k$ $P_{k\|-} = P_{k-1} + Q$	$K_k = P_{k\|-}\left(P_{k\|-} + R\right)^{-1}$ $\hat{x}_k = \hat{x}_{k\|-} + K_k\left(z_k - \hat{x}_{k\|-}\right)$ $P_k = (I - K_k)P_{k\|-}$

Some definitions:

- K_k is the Kalman gain;
- P_k is the state covariance matrix

The table below shows the process for the first 6 measurements, with the assumptions that $P_0 = 100$, $R = 1$, $Q = 0.01$, and $x_0 = 0$.

time, k (sec)	zk (measurements)	vk (measurement noise)	qk (process noise)	Pk (Covariance)	Xk (Estimate)	Pk-minus	Xk-minus	Kk (Kalman Gain)
0	10.67	1	0.1	100.000	0.000	100.100	0.000	0.990
5	26.79	1	0.1	0.989	26.523	1.089	26.523	0.521
10	16.66	1	0.1	0.473	21.380	0.573	21.380	0.364
15	10.32	1	0.1	0.301	17.350	0.401	17.350	0.286
20	29.98	1	0.1	0.215	20.964	0.315	20.964	0.239
25	21.70	1	0.1	0.163	21.139	0.263	21.139	0.208

A plot of the state estimate of the signal measurements is provided in Figure 8. Note that the estimate follows a nominally mean path relatively unaffected by the spikiness of the measurements. This is not to imply that this result is a measure of "goodness": the lack of responsiveness can translate into incorrectly masking real events.

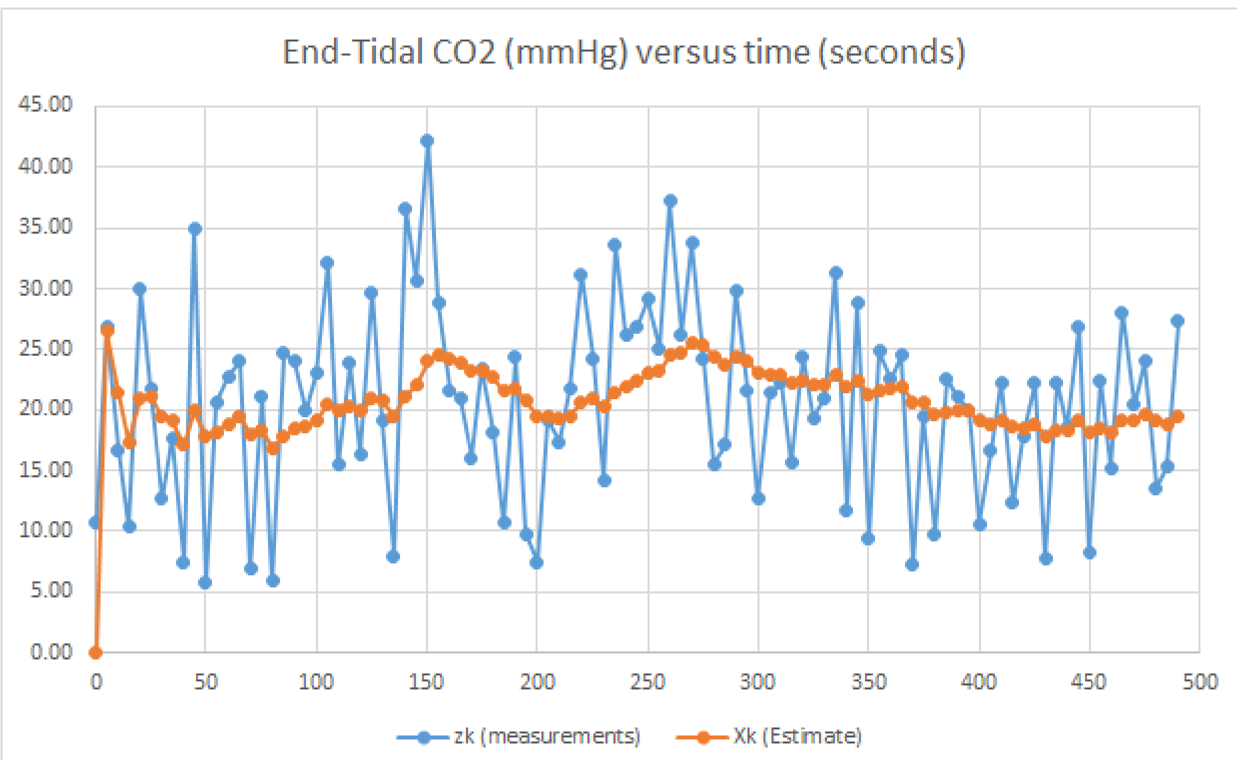

FIGURE 8: KALMAN FILTER TRACKING OF END-TIDAL CO2 WITH MEASUREMENT ERROR SET TO 1 MMHG.

Indeed, if we look at the number of occurrences of consecutive signals of 3 measurements at or below 15 mmHg, the following graph, Figure 9, is quite telling in terms of the responsiveness of the state estimates relative to the actual raw data.

Compared with the raw data, there are no reports of state estimates with consecutive measurements below the emergent threshold of 15 mmHg.

Now, compare this result with that of Figures 10 and 11, in which the measurement error has been reduced to 0.1 mmHg. The response of the state estimate follows more closely to the raw data measurement, and, consequently, the reports of simultaneous measurements dropping below the emergent threshold have also increased.

Indeed, comparing Figure 9 with Figure 11 shows much more responsiveness, albeit not at the same level as the raw measurements. When we further reduce the measurement error, first to 0.01 (Figures 12 and 13), the number of simultaneous measurements increases corresponding to the raw data. When the measurement error is effectively made close to zero (error = 0.001, Figures 14 and 15), the state estimate is identical to the measurement, and all occurrence of events falling below threshold are reported.

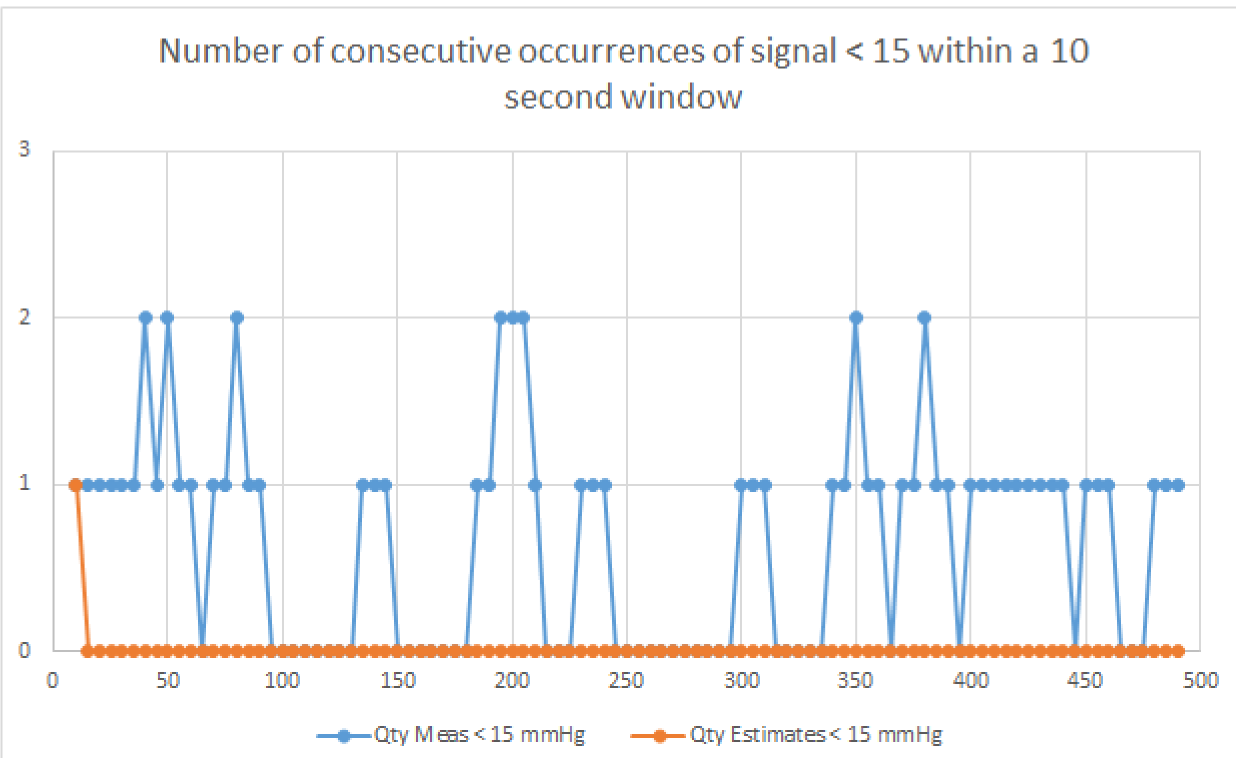

FIGURE 9: NUMBER OF OCCURRENCES OF MEASUREMENTS CONSECUTIVELY IN WHICH THE MEASUREMENTS DROP BELOW 15 MMHG WITH MEASUREMENT NOISE AT 1 MMHG.

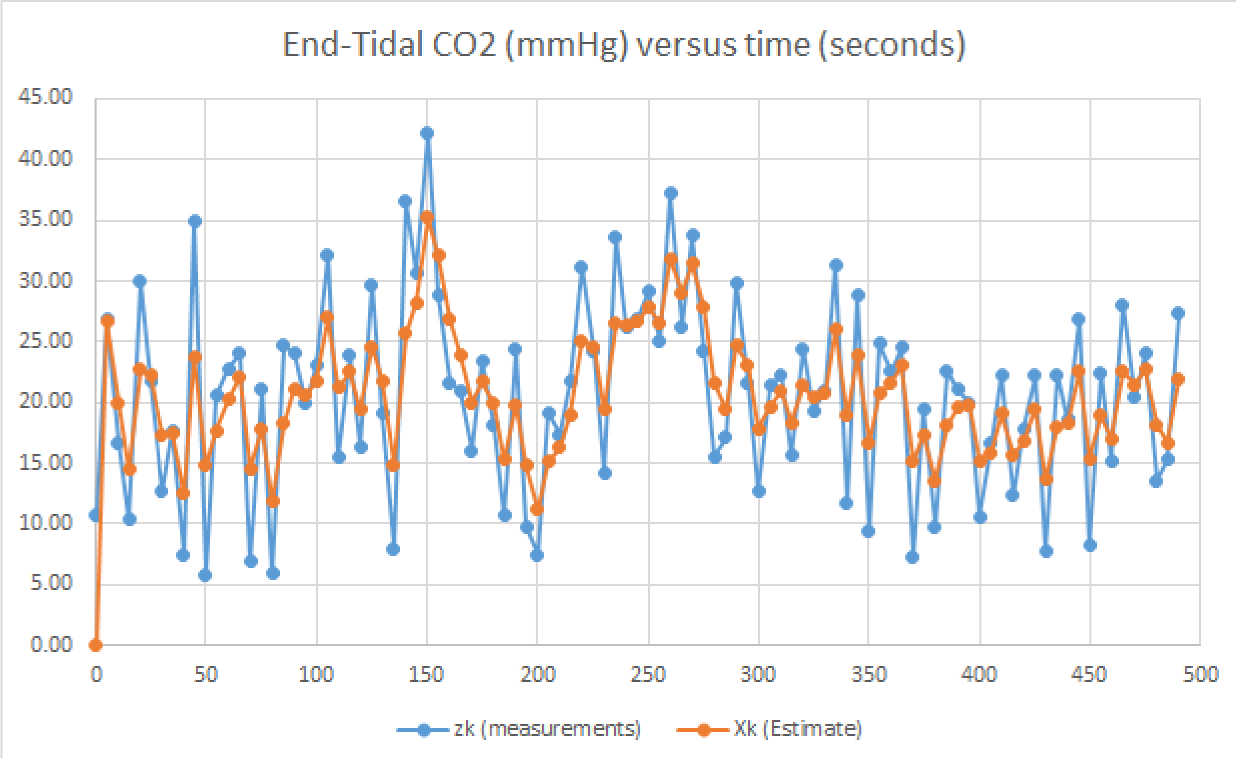

FIGURE 10: KALMAN FILTER TRACKING OF END-TIDAL CO_2 WITH MEASUREMENT ERROR SET TO 0.1 MMHG.

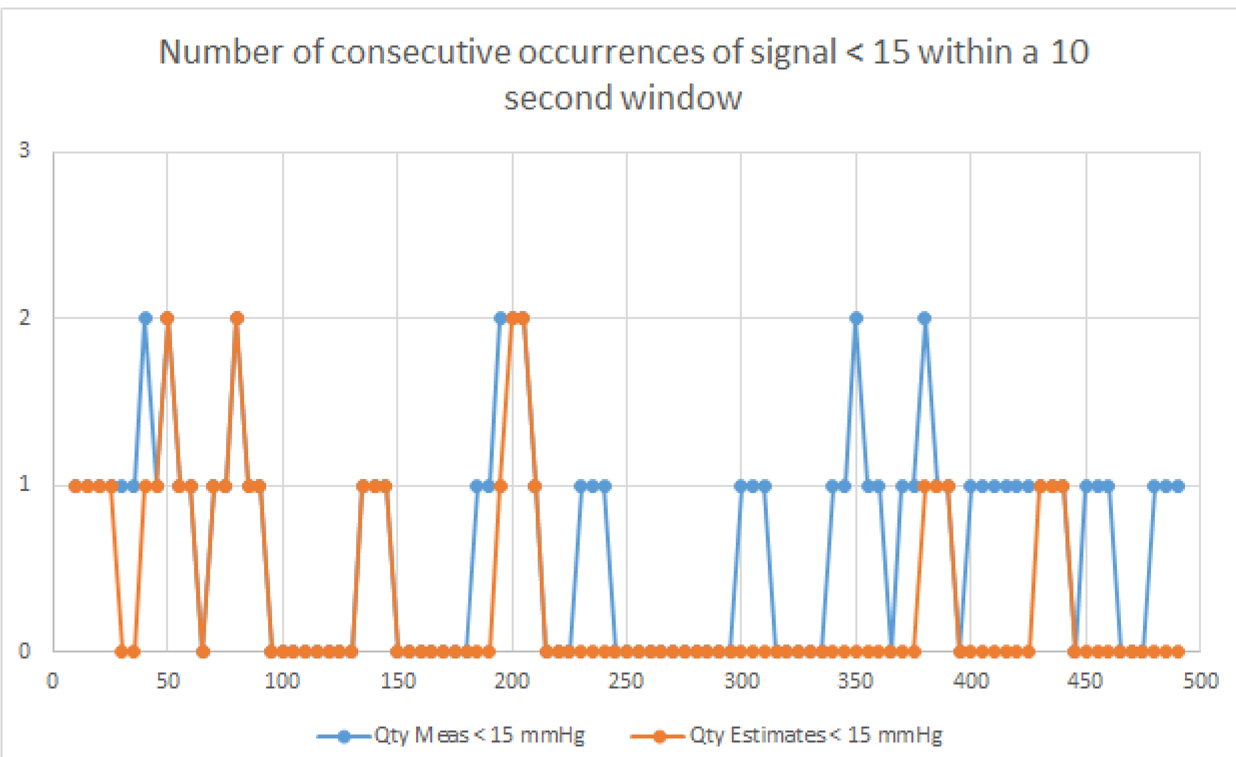

FIGURE 11: NUMBER OF OCCURRENCES OF MEASUREMENTS CONSECUTIVELY IN WHICH THE MEASUREMENTS DROP BELOW 15 MMHG WITH MEASUREMENT NOISE AT 0.1 MMHG.

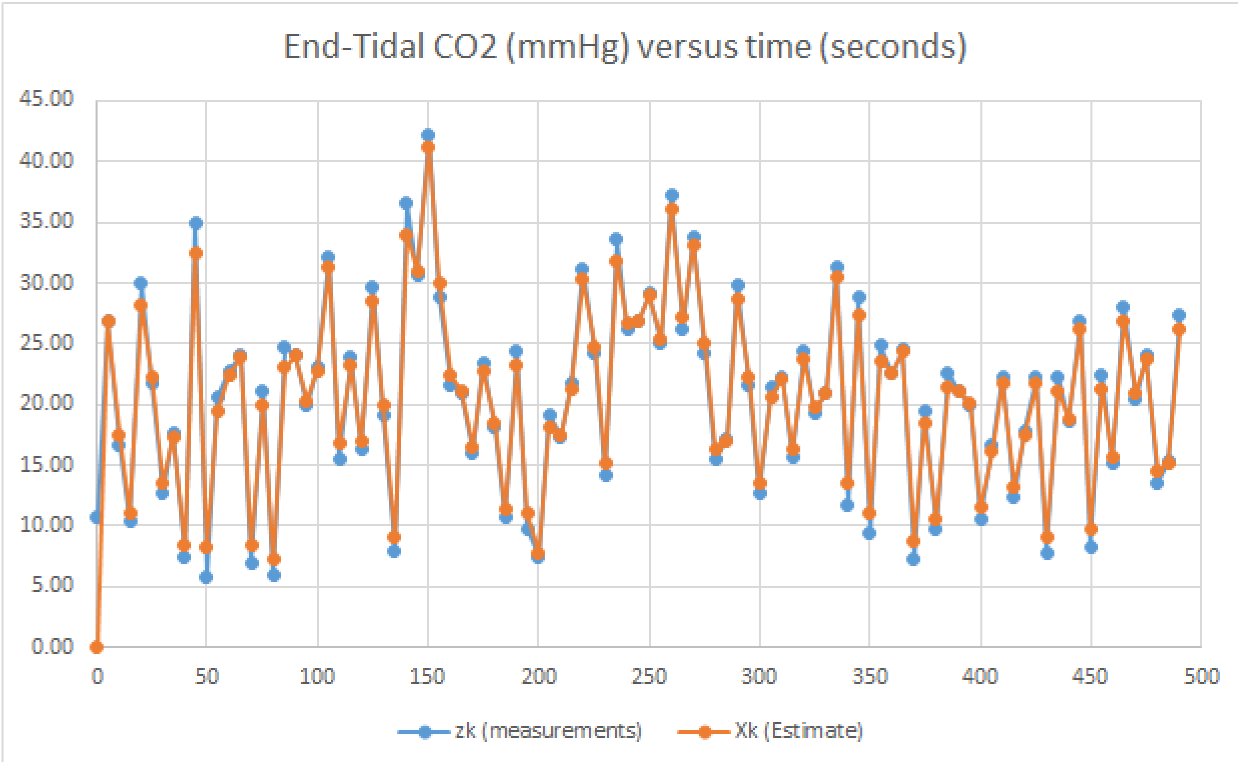

FIGURE 12: KALMAN FILTER TRACKING OF END-TIDAL CO2 WITH MEASUREMENT ERROR SET TO 0.01 MMHG.

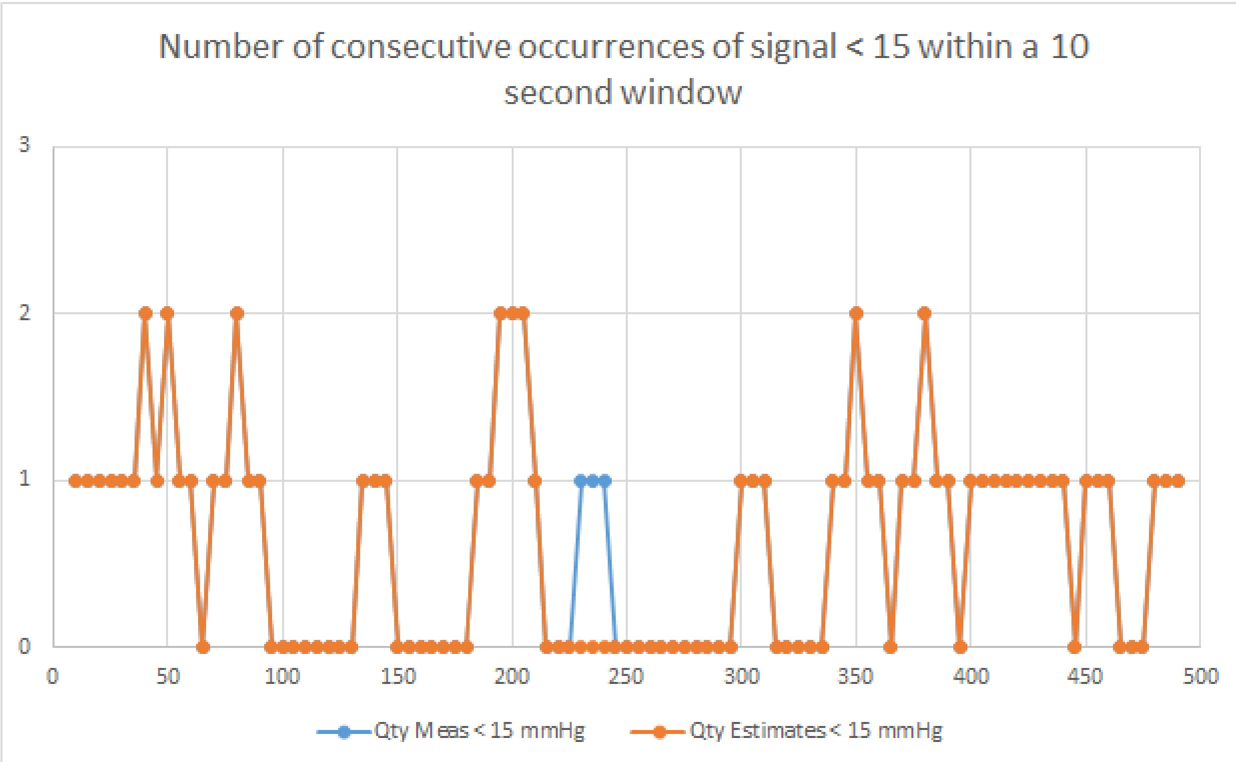

FIGURE 13: NUMBER OF OCCURRENCES OF MEASUREMENTS CONSECUTIVELY IN WHICH THE MEASUREMENTS DROP BELOW 15 MMHG WITH MEASUREMENT NOISE AT 0.01 MMHG.

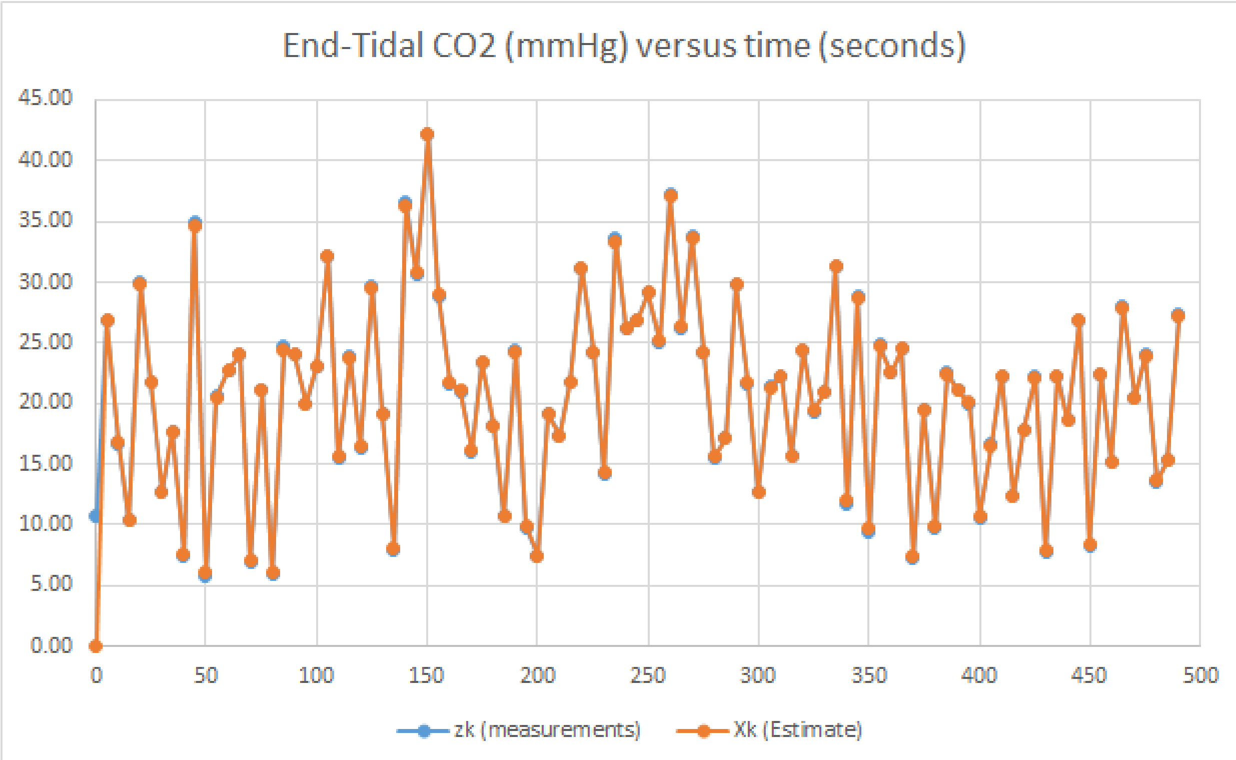

FIGURE 14: KALMAN FILTER TRACKING OF END-TIDAL CO2 WITH MEASUREMENT ERROR SET TO 0.001 MMHG.

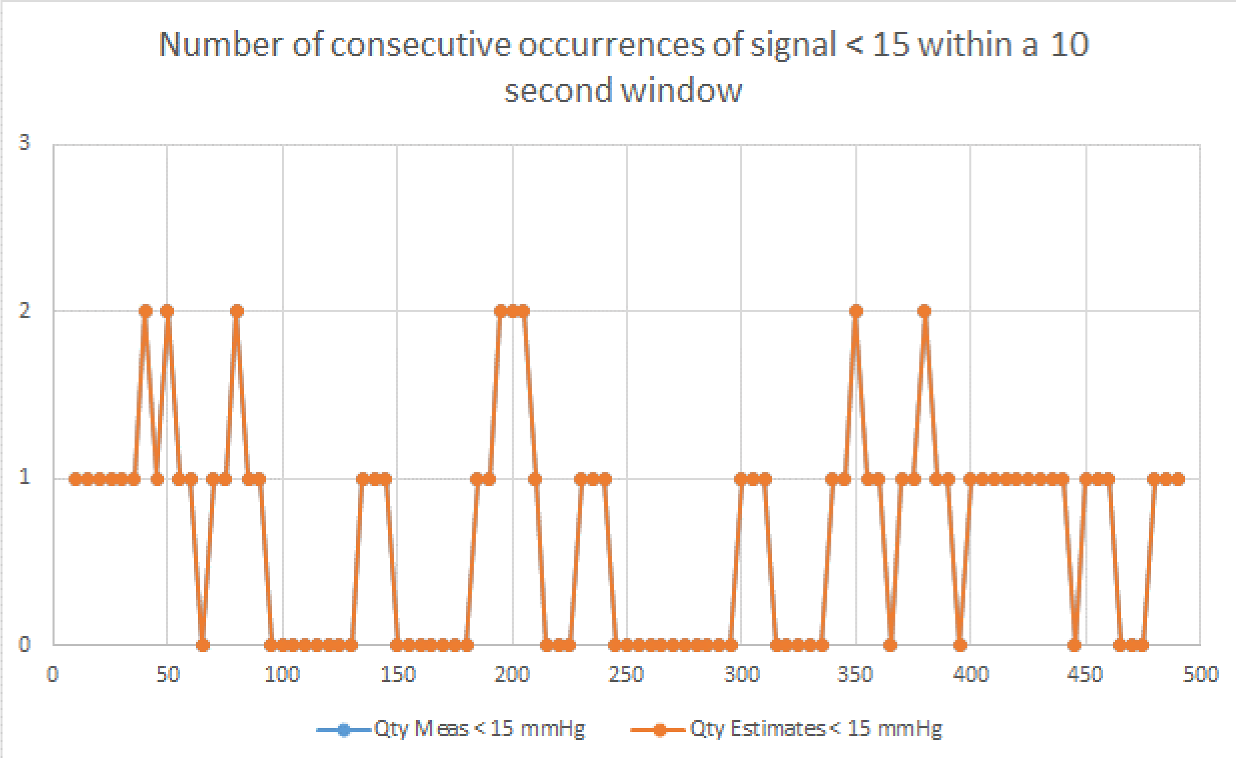

FIGURE 15: NUMBER OF OCCURRENCES OF MEASUREMENTS CONSECUTIVELY IN WHICH THE MEASUREMENTS DROP BELOW 15 MMHG WITH MEASUREMENT NOISE AT 0.001 MMHG.

Observations

If the objective is to minimize the number of false alarms, and to define a true event as the multiple, repeated occurrence of a measurement, it may be possible to tailor such thresholds by intelligently selecting the sensitivity of the model for the raw signal data. A review of Figure 11 seems to show that a balance between number of consecutive occurrences and minimizing individual occurrence false alarm rate is being approached (not achieved quite fully, but approached).

The selection of thresholds must be within the purview and control of the licensed clinician as part of the practice of medicine. Technology cannot make these decisions as technical algorithms would need to take into account the full context of the patient and the training of the end user. Maybe someday this will be possible (if desirable), but it is certainly not the case today.

Nonetheless, the objective was to present the sensitivity analysis and demonstrate several models to trigger further discussion and investigation in the mind of the researcher and the clinician.

References

[1] Ilene MacDonald, "Hospitals rank alarm fatigue as top patient safety concern", Fierce Healthcare. January 22, 2014.

[2] Wong, Michael; Mabuyi, Anuj; Gonzalez, Beverly; "First National Survey of Patient-Controlled Analgesia Practices." March-April 2013, A Promise to Amanda Foundation and the Physician-Patient Alliance for Health & Safety.

[3] Ibid.

[4] Ibid.

[5] Ibid.

[6] Souto, R.F.; Ishihara, J.Y.; Borges, G. A., "A Robust Extended Kalman Filter for Discrete-time Systems with Uncertain Dynamics, Measurements and Correlated Noise." 2009 American Control Conference. Hyatt Regency Riverfront, St. Louis, MO. June 10-12 2009.

[7] Sameni, R.; Shamsollahi, M.B.; Jutten, C., "Filtering Electrocardiogram Signals Using the Extended Kalman Filter." Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China, September 1-4, 2005.

[8] Lindsay Kleeman, "Understanding and Applying Kalman Filtering." Teaching Notes. Lindsay.Kleeman@monash.edu.au.