Chapter 9

Interesting Results

Going a little bit further than what the goal of this thesis pretends, some interesting tools discovered during the development of this work are briefly explained in this chapter.

9.1 The sample variance in the presence of steps

As it has been said, expression $E(S^2) = \sigma$ is true just for stationary processes. This makes it unappropriate for estimating the standard deviation under non-stationary processes, not only when there are vibrations, but also right after a leap or step in the reading.

The evolution over time of $E(S^2)$ after a step is mathematically studied in appendix A. Results show:

$$E(S^2) = \sigma_{proc}^2 + \frac{(N-t)t}{N} \Delta\mu^2. \tag{9.1}$$

Where σ_{proc}^2 is the variance of the underlying stationary process and $\Delta\mu$ is the size of the step, or the increment in the mean of the process, t is the elapsed time after the leap measured in samples and N is the sample size.

This equation corresponds to a parabola that starts at $t = 0$ and dies at $t = N$ with expected value σ_{proc}^2. Its expected maximum is $E(S^2)|_{max} = \frac{1}{4}\frac{N}{(N-1)}\Delta\mu^2$ and takes place at $t = \frac{N}{2}$

Figure 9.1 shows the evolution of S^2 calculated over real data. In this case predicted peaks are very clear and distinct.

Although results are not always so clear as on figure 9.1 due to the uncertainty of the estimation, the double-queue algorithm explained in chapter 8 can be used as well in this situation obtaining more reliable values of the sample variance. The double queue algorithm is the easiest low-pass filter we can implement and, as figure 9.2 shows, it makes the peaks appear clearly in graphics.

So, statistic S^2 is found to be a powerful and appropriate *"events"* detector. Peaks on the reading of S^2 provide information, not only about weight changes, but also about how sharply these changes happen. Thereby, it could be suitable for detecting small instant increments of the weight in the scale. The monitor-

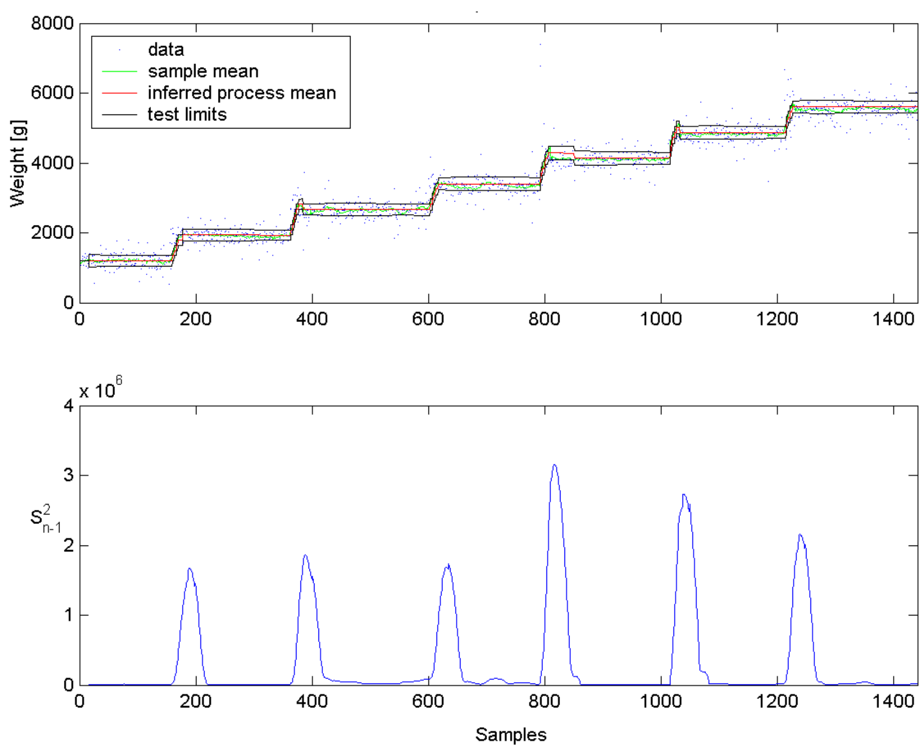

Figure 9.1: Evolution of S^2 shows peaks over real data as predicted.

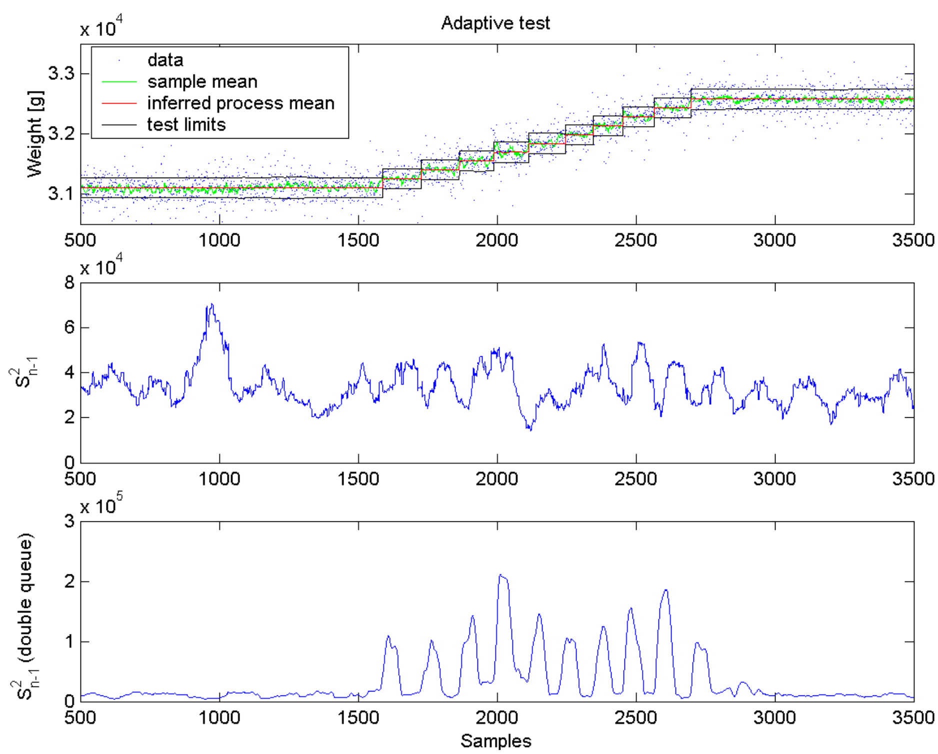

Figure 9.2: Evolution of S^2 during small noisy steps calculated directly and with the double-queue algorithm.

ing sample by sample (not only once every N samples) of this statistic ensures its evaluation at the peak, which could be used to faster determine the weight increment.

In the same way, it is a helpful tool for segmenting data series during the initial parameters estimation of the process as much for the mean value as for the variance. By segmenting data, they can be analyzed separately obtaining more accurate results in an automatic, faster and easier way not even having to worry that the process is stationary or that units are not falling into the scale.

9.2 A *non-stationarity* estimator. The statistic $S^2 - S_{\bar{x}}^2 N$

Vibrations of the ground make measurements spread and, consequently, they affect the sample standard deviation, whose expected value is no longer equal to σ_{proc} ($E(S^2) \neq \sigma_{proc}^2$). Appendix B contains a mathematical exposition of the expected value of the sample mean and the sample variance in the case of induced vibrations. Results show that equation 9.2 is *no longer* valid.

$$E(S_{proc}^2) = E(S_{\bar{x}}^2)N. \tag{9.2}$$

However, that equation is not useless by any means. On the contrary, this fact suggests that the statistic $S_{proc}^2 - S_{\bar{x}}^2 N$, whose expected value in a stationary situation is zero, would be a good estimator of non-stationarity.

$$NSE = E(S_{proc}^2) - E(S_{\bar{x}}^2)N. \tag{9.3}$$

Its behavior and expected value when the input is superposed to a sine-wave is studied in appendix B. But this statistic is affected by every kind of non-stationarities, not just by oscillations, but also by steps. Thereby, its evolution right after a step has been studied as well. Fortunately it is easy since we already know equation (9.1), which applied to the sample mean provides:

$$E(S_{\bar{x}}^2) = \sigma_{\bar{x}}^2 + \frac{(N-t)t}{N}\Delta\mu^2. \tag{9.4}$$

Since $\sigma_{\bar{x}}^2 N = \sigma_{proc}^2$, substitution of equations (9.1) and (9.4) in the definition of the non-stationarity estimator (NSE) (equation 9.3) leads to:

$$E(NSE) = \frac{(N-t)t}{N}\Delta\mu^2(1-N). \tag{9.5}$$

That represents a parabola with a minimum at $t = \frac{N}{2}$ with value $E(NSE)\|_{min} = -\frac{N}{4}\Delta\mu^2$.

Figure 9.3 presents the evolution of the NSE when aplied to the same non-stationary data set we used previously. The graphic proves how clearly the NSE alerts about the presence of vibrations. Inverted peaks are found too as predicted. Let us note that negative peaks are always preceded by small positive peaks and, when steps are small, there is even no negative peak, but just the positive one. This happens due to the different sample sizes used for the estimation of terms S^2

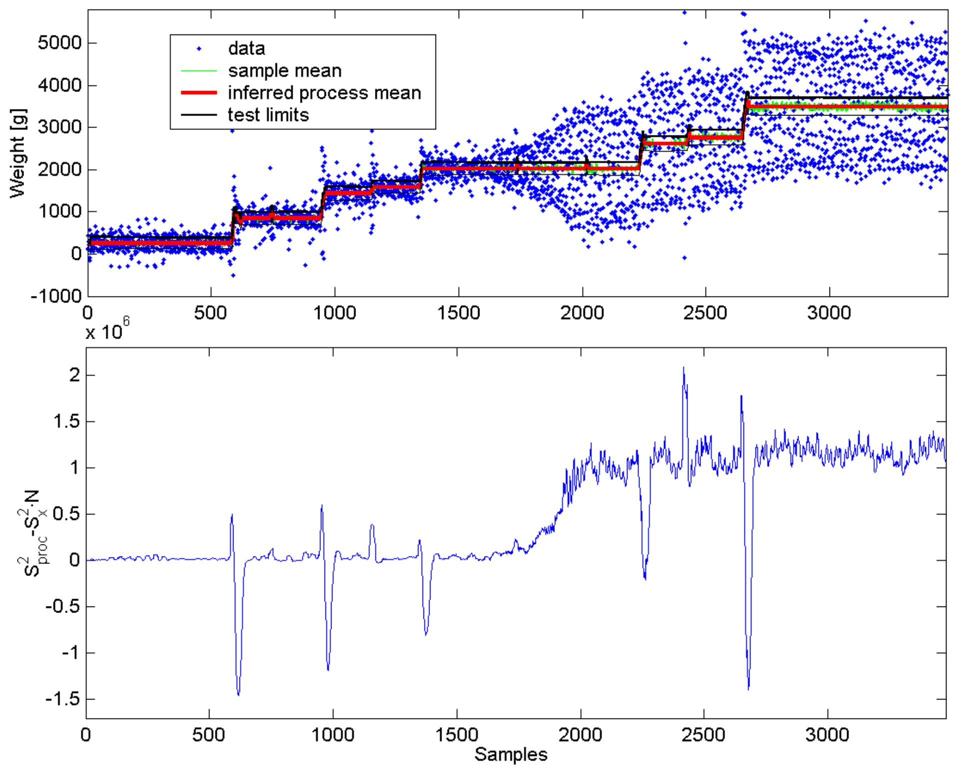

Figure 9.3: Evolution of the NSE during transition from stationary to non-stationary process.

(smaller) and $S^2_{\bar{x}}$ (bigger). That makes the first one grow earlier than the second one, so the positive peak predominates for a while. This could be avoided (if desired) by on purpose introducing a delay time for S^2 or by adapting the sample sizes so that they are similar.

This statistic can be used to select data subsets corresponding to stationary situations out from a given long data set. This helps in the initial estimation of the process parameters, specially for the variance of the process, whose estimation is strongly altered by non-stationarities. However, since selected data could be still be affected by low oscillation levels, the double queue algorithm should be used in this application to improve the accuracy as it has proved to do.